

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7: C12N 9/00 A2		(11) International Publication Number: WO 00/50575		
		(43) International Publication Date: 31 August 2000 (31.08.00)		
(21) International Application Number: PCT/US (22) International Filing Date: 25 February 2000 ((81) Designated States: CA, CN, JP, MX, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).			
(30) Priority Data: 60/121,968 26 February 1999 (26.02.99) t	Published Without international search report and to be republished upon receipt of that report.		
(71) Applicant: CALGENE LLC [US/US]; 1920 Fifth Stre CA 95616 (US).	et, Dav	is,		
(72) Inventor: MATSUI, Kenji; 2-51, Nishiki-cho, Y City, Yamaguchi 753-0068 (JP).	amaguo	hi		
(74) Agents: SCHWEDLER, Carl, J. et al.; Calgene L. Fifth Street, Davis, CA 95616 (US).	LC, 19	20		
(54) Title: EATTY ACID O HYDDODEDOVIDE I VAS				

(54) Title: FATTY ACID 9-HYDROPEROXIDE LYASE NUCLEIC ACID SEQUENCES

(57) Abstract

This invention relates to 9-Hydroperoxide Lyase or 9-HPO lyase enzymes. DNA constructs useful for the expression of a plant HPO lyase in a cell are described. Furthermore, DNA constructs useful for the antisense expression of a 9-HPO lyase in a plant cell are described. Such constructs will contain a DNA sequence encoding the 9-HPO lyase of interest under the control of regulatory elements capable of preferentially directing the expression of the 9-HPO lyase in plant tissue, when such a construct is expressed in a trangenic plant. This invention also relates to methods of using a DNA sequence encoding a 9-HPO lyase for the modification of the volatile aldehydes in plant tissues, as well as for methods of increasing disease resistance in a plant. 9-HPO Lyase sequences exemplified herein are obtained from *Arabidopsis*.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCI

AL AM AT AU AZ BA BB BE BF BG BJ BR CCF CG CH CI CM CU CZ DE DK EE	Albania Armenia Armenia Austria Australia Azerbaijan Bosnia and Herzegovina Barbados Belgium Burkina Faso Bulgaria Benin Brazil Belarus Canada Central African Republic Congo Switzerland Côte d'Ivoire Cameroon China Cuba Czech Republic Germany Denmark Estonia	ES FI FR GA GB GE GN GR HU IE IL IS IT JP KE KG KP KZ LC LL LK LR	Spain Finland France Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland Israel Iceland Italy Japan Kenya Kyrgyzstan Democratic People's Republic of Korea Republic of Korea Kazakstan Saint Lucia Liechtenstein Sri Lanka Liberia	LS LT LU LV MC MD MG MK ML MN MR MW MX NE NL NO NZ PL PT RO RU SD SE SG	Lesotho Lithuania Luxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali Mongolia Mauritania Malawi Mexico Niger Netherlands Norway New Zealand Poland Portugal Romania Russian Federation Sudan Sweden Singapore	SI SK SN SZ TD TG TJ TM TR TT UA UG US UZ VN YU ZW	Slovenia Slovakia Senegal Swaziland Chad Togo Tajikistan Turkmenistan Turkey Trinidad and Tobago Ukraine Uganda United States of America Uzbekistan Viet Nam Yugoslavia Zimbabwe
--------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Fatty Acid 9-Hydroperoxide Lyase Nucleic Acid Sequences

5

10

25

30

35

INTRODUCTION

Technical Field

This invention relates to the application of genetic engineering techniques to plants. More specifically, the invention relates to plant hydroperoxide lyase sequences and methods for the use of such sequences.

Background

With the development of genetic engineering techniques, it is now possible to transfer genes from a variety of organisms into the genome of a large number of different plant species. This process has many advantages over plant breeding techniques, as genes may now be transferred from one plant species to another plant species, rather than simply from a plant to the same, or different, but closely related, species.

Degradation of polyunsaturated fatty acids starts by the oxygenation at cis-cis double bonds of polyunsaturated fatty acids. This reaction is catalyzed by lipoxygenase (EC 1.13.11.12) enzymes which are present in plants, animals and microorganisms. The oxygenated products, called fatty acid hydroperoxides, are precursors for many important hormones (e.g. lipoxins, jasmonic acid, traumatic acid) and flavor/fragrance molecules (e.g. cis-3-hexenol, 1-octen-3-ol) in plants.

Compounds, such as jasmonic acid, are produced from hydroperoxides, such as 13-hydroperoxylinolenic acid, via an allene oxide synthase (referred to as AOS) and an allene oxide cyclase (referred to as ACS)-dependent pathway. Jasmonic acid is involved in stress and disease resistance signaling responses via the octadecanoid pathway. 13-hydroperoxylinolenic can also be catabolized by peroxygenases to form cutin monomers. Alternatively, 13-hydroperoxylinolenic

can be catabolyzed by hydroperoxide lyase eventually forming volatile aldehydes and traumatic acid.

Fatty acid hydroperoxide lyase (HPO lyase) catalyzes the cleavage of carbon-carbon bonds in polyunsaturated fatty acid hydroperoxides to produce short-chain aldehydes and ω -oxoacids (Vick, et al. (1976) Plant Physiol. 57:780-788). products of lysis of fatty acid hydroperoxides, such as shortchain volatile aldehydes are common in plant species. aromas/flavors produced are a function of the specific 10 products produced by the lysis of specific fatty acid hydroperoxides. For example, some short-chain volatile aldehydes contribute to the "green notes" in a wide variety of plant leaves, vegetables and fruits. "Green notes" are volatile molecules that contribute to the organoleptic 15 qualities of flavor and fragrance of edible plant tissues. These qualities are often referred to as grassy, or "green" characteristics. Other short-chain volatile aldehydes, such as (3Z, 6Z)-nonadienol produced by the lysis of fatty acid 9hydroperoxide by a fatty acid 9-hydroperoxide lyase (9-HPO 20 lyase or 9-HPOL), contribute a melon aroma and/or a melon flavor, or sometimes referred to as a "melon" or "fresh" characteristic, to fruits and vegetables. characteristics are important to industries concerned with fragrances and flavorings.

Furthermore, short-chain aldehydes are also thought to be involved in disease resistance. For example, Croft, et al ((1993) Plant Physiol. 101:13-24) recently reported that (3Z)-hexenol and (2E)-hexenal levels increased during a hypersensitive-response in kidney bean plants. In addition, they also demonstrated that (2E)-hexenal is an effective antibacterial agent.

The characterization of hydroperoxide (also referred to as HPO lyase or HPOL) is useful for the further study of plant fatty acid metabolism systems and for the development of transgenic plant with increased organoleptic properties, including aromas and flavors. Studies of plant mechanisms may provide means to further enhance, control, modify, or otherwise alter the organoleptic qualities of edible plant

tissues. Furthermore, the elucidation of the physiological roles of HPO lyase and it's products may be useful for the further study of disease resistance responses, such as the HR response. Of particular interest are the nucleic acid sequences of genes encoding proteins which may be useful for applications in genetic engineering.

Relevant Literature

5

10

15

20

30

35

Shibata, et al. (1995) Plant Cell Physiol 36:147-156 reports the purification of a HPO lyase from green bell pepper. Matsui, et al. (1996) FEBS Letters 394:21-24 reports the cloning of a HPO lyase from bell pepper and the protein encoded by the HPO lyase sequence resembles a cytochrome P450.

SUMMARY OF THE INVENTION

The present invention provides nucleic acid sequences encoding for proteins which catalyze the formation of short-chain aldehydes and oxo-acids from 9-hydroperoxide fatty acids. Such proteins are referred to herein as 9-hydroperoxide lyases or 9-HPO lyase.

By this invention, nucleic acid sequences encoding plant 9-HPO lyase may now be characterized with respect to enzyme activity.

Thus, this invention encompasses 9-HPO lyase nucleic acid sequences and the corresponding amino acid sequences, and the use of these nucleic acid sequences in the preparation of oligonucleotides containing 9-HPO lyase encoding sequences for analysis and recovery of plant 9-HPO lyase gene sequences. The 9-HPO lyase encoding sequence may encode a complete or partial sequence depending upon the intended use. All or a portion of the genomic sequence, or cDNA sequence, is intended.

Of special interest are recombinant DNA constructs which provide for transcription or transcription and translation (expression) of the 9-HPO lyase sequence in a host cell. In particular, constructs which are capable of transcription or transcription and translation in plant host cells are preferred. For some applications a reduction in 9-HPO lyase

may be desired. Thus, recombinant constructs may be designed having the plant 9-HPO lyase sequences in a reverse orientation for expression of an anti-sense sequence or use of co-suppression, also known as "transwitch", constructs may be useful. Such constructs may contain a variety of regulatory regions including transcriptional initiation regions obtained from genes preferentially expressed in plant seed tissue. For some uses, it may be desired to use the transcriptional and translational initiation regions of the 9-HPO lyase gene either with the 9-HPO lyase encoding sequence or to direct the transcription and translation of a heterologous sequence.

In yet a different aspect, this invention relates to a method for producing a 9-HPO lyase in a host cell or progeny thereof via the expression of a construct in the cell. Cells containing the 9-HPO lyase as a result of the production of the 9-HPO lyase encoding sequence are also contemplated herein.

In addition, methods for modifying the volatile aldehyde content in plant tissues, especially fruit and leaf tissue, as well as methods for producing transgenic plants with increased disease resistance are contemplated.

Also considered in this invention are the modified plants, tissues and aldehyde compositions obtained by expression of the 9-HPO lyase sequence and protein of this invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the complete nucleotide sequence of the Arabidopsis HPO lyase.

Figure 2 shows a comparison of the amino acid sequences of the bell pepper HPO lyase and the *Arabidopsis* HPO lyase-like sequence.

Figure 3 shows a comparison of the amino acid sequences of the Arabidopsis allene oxide synthase and the Arabidopsis HPO lyase-like sequence.

Figure 4 shows the complete nucleotide sequence of the tomato HPO lyase.

35

5

10

15

20

25

Arter Carlos and Carlos

 $\label{eq:continuous_problem} \mathcal{C}(\mathbf{r}) = (\mathbf{r} - \mathbf{r}) \cdot (\mathbf{r} + \mathbf{r}) + (\mathbf{r} + \mathbf{r}) \cdot (\mathbf{r} + \mathbf{r}) \cdot (\mathbf{r} + \mathbf{r}) + (\mathbf{r} + \mathbf{r}) \cdot (\mathbf{r} + \mathbf{r}) \cdot (\mathbf{r} + \mathbf{r}) + (\mathbf{r} + \mathbf{r}) \cdot ($

Figure 5 shows the complete nucleotide sequence of the cucumber allene oxide synthase.

Figure 6 shows the complete nucleotide sequence of the cucumber 9-Hydroperoxide Lyase.

Figure 7 shows the amino acid sequence alignment between the bell pepper, banana, and *Arabidopsis* HPO lyase, with the highly conserved peptide sequences highlighted.

Figure 8 Provides the percent similarity in the upper right corner and the percent divergence in the lower right corner for the nucleotide sequences (Figure 8A) and amino acid sequences (Figure 8A) of the bell pepper HPOL (CaHPOL), tomato fruit HPOL (LeHPOL), cucumber hypocotyl HPOL (CsC17HPOL, pseudogene), Arabidopsis inflorecence HPOL (AtHPOL), banana leaf HPOL (MsHPOL), cucumber hypocotyl 9-HPOL (Cs15HPOL), Guayule AOS (GuAOS), flaxseed AOS (LiAOS), and the Arabidopsis AOS (AtAOS).

Figure 9 shows the gas chromatography (GC) analysis of the cucumber 9-HPO lyase using linoleic acid 13-hydroperoxide (Figure 9A) and linoleic acid 9- hydroperoxide (figure 9B) substrates.

Figure 10 provides the results of the spectrophotometric assay of the cucumber 9-HPO lyase expressed from *E. coli* using linoleic acid 13-hydroperoxide and linoleic acid 9-hydroperoxide substrates.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the subject invention, nucleotide sequences are provided which are capable of coding sequences of amino acids, such as, a protein, polypeptide or peptide, which demonstrate the ability to form short-chain aldehydes and oxo-acids from 9-hydroperoxide fatty acids under plant enzyme reactive conditions. By "enzyme reactive conditions" is meant that any necessary conditions are available in an environment (i.e., such factors as temperature, pH, lack of inhibiting substances) which will permit the enzyme to function.

In one embodiment of the present invention, nucleic acid sequences are provided which encode for 9-hydroperoxide lyase

5

10

15

20

25

.30

WO 00/50575 6 PCT/US00/05311

(hereinafter referred to as 9-HPO lyase). Such nucleic acid sequences are demonstrated herein to encode enzymes which form (3Z,6Z)-nonadienal and 9-oxo-nonanoic acid from Linolenic acid 9-hydroperoxide or 9-hydroperoxy-(10E, 12Z, 15Z)-octadecadienoic acid.

In the examples provided below, a nucleic acid sequence from cucumber (*Cucumis sativus*) is identified from cDNA libraries made from total RNA isolated from cucumber hypocotyls. A full length coding sequence is obtained, and the product encoded by the full length sequence demonstrates activity towards the substrate linolenic acid 9-hydroperoxide to produce (3Z, 6Z)-nonadienal and 9-oxo-nonanoic acid.

In another embodiment of the present invention, expression constructs are provided which direct the expression of nucleic acid sequences encoding 9-HPO lyase in bacterial and plant tissues.

Of particular interest in the present invention, is the use of such expression constructs to produce transgenic plants with increased production of short-chain volatile aldehydes in plant fruits and tissues. Such volatile aldehydes are important constituents of the characteristic flavors of fruits, vegetables and green leaves. Thus, the 9-HPO lyase sequence of the present invention may be used in expression constructs to produce transgenic plants with improved organoleptic properties, in particular improved fruity (or "melon") note flavor and aroma characteristics.

The nucleic acid sequences of the present invention may also find use in expression constructs for the production of transgenic plants with increased resistance to various pathogens. Transgenic plants expressing the HPO lyase sequence of the present invention may exhibit an enhanced hypersensitive-reaction (HR response) in response to pathogen attack due to the increased production of aldehydes involved in the HR response, such as (3Z)-hexenal and (2E)-hexenal (Croft, et al. (1993) Plant Physiol. 101:13-24). Aldehydes, such as (2E)-hexenal, have also been shown to be effective anti-bacterial agents, further contributing to enhanced disease resistance (Croft, et al. (1993), supra).

5

10

15

20

25

30

Furthermore, these compounds may be involved in a general wounding response in plants.

Also of particular interest in the present invention is the use of 9-HPO lyase nucleic acid sequences in constructs to direct the expression of 9-HPO lyase in a prokaryotic and/or eukaryotic host cells for the production of flavorings and aromas.

To confirm the activity and specificity of proteins encoded by nucleic acid sequences as 9-HPO lyase enzymes, assays are performed on protein extracts of cultured bacterial cells expressing the protein. As described in more detail in the examples, cucumber 9-HPO lyase expression constructs are prepared to direct the expression of the 9-HPO lyase sequence in *E coli*. The expressed 9-HPO lyase enzymes demonstrate the ability to produce (3Z, 6Z)-nonoadienal and 9-oxo-nonanoic acid from linolenic acid 9-hydroperoxide or 9-hydroperoxy-(10E, 12Z, 15Z)-octadecadienoic acid by gc analysis.

The skilled artisan will recognize that several methods for the identification of additional sequences encoding 9-HPO lyase are available in the art.

For example, antibody preparations, nucleic acid probes (DNA and RNA) and the like may be prepared and used to screen and recover "homologous" or "related" HPO lyase sequences from a variety of plant sources. Typically, nucleic acid probes are labeled to allow detection, preferably with radioactivity although enzymes or other methods may also be used. For immunological screening methods, antibody preparations either monoclonal or polyclonal may be utilized. Polyclonal antibodies, although less specific, typically are more useful in gene isolation. For detection, the antibody is labeled using radioactivity or any one of a variety of second antibody/enzyme conjugate systems that are commercially available. Examples of some of the available antibody detection systems are described by Oberfilder (Focus (1989) BRL Life Technol., Inc., 11:1-5).

In order to obtain additional HPO lyase sequences, a genomic or other appropriate library prepared from the candidate plant source of interest may be probed with

5

10

15

20

25

30

conserved sequences from one or more plant HPO lyase(s) to identify homologously related sequences. Positive clones may be analyzed by restriction enzyme digestion and/or sequencing. When a genomic library is used, one or more sequences may be identified providing both the coding region, as well as the transcriptional regulatory elements of the HPO lyase gene from such plant source. Probes can also be considerably shorter than the entire sequence. Oligonucleotides may be used, for example, but should be at least about 10, preferably at least about 15, more preferably at least 20 nucleotides in length. When shorter length regions are used for comparison, a higher degree of sequence identity is required than for longer sequences. Shorter probes are often particularly useful for polymerase chain reactions (PCR), especially when highly conserved sequences can be identified. (See, Gould, et al., PNAS USA (1989) 86:1934-1938.)

When longer nucleic acid fragments are employed (>100 bp) as probes, especially when using complete or large cDNA sequences, one can still screen with moderately high stringencies (for example using 50% formamide at 37°C with minimal washing) in order to obtain signal from the target sample with 20-50% deviation, i.e., homologous sequences. (For additional information regarding screening techniques see Beltz, et al., Meth. Enzymology (1983) 100:266-285).

Homologous sequences are found when there is an identity of sequence and may be determined upon comparison of sequence information, nucleic acid or amino acid, or through hybridization reactions between a known HPO lyase and a candidate source. Conservative changes, such as Glu/Asp, Val/Ile, Ser/Thr, Arg/Lys and Gln/Asn may also be considered in determining sequence homology. Typically, a lengthy nucleic acid sequence may show as little as 50-60% sequence identity, and more preferably at least about 70% sequence identity, between the target sequence and the given plant HPO lyase of interest excluding any deletions which may be present, and still be considered related. Amino acid sequences are considered homologous by as little as 20% sequence identity between the two complete mature proteins.

5

10

15

20

25

30

(See generally, Doolittle, R.F., OF URFS and ORFS (University Science Books, CA, 1986.)

In addition, not only can sequences provided herein be used to identify homologous 9-hydroperoxide lyases, but the resulting sequences obtained therefrom may also provide a further method to obtain plant 9-hydroperoxide lyase sequences from other plant sources. In particular, PCR may be a useful technique to obtain related 9-HPO lyases from sequence data provided herein. One skilled in the art will be able to design oligonucleotide probes based upon sequence comparisons or regions of typically highly conserved sequence. Furthermore, additional sequences may be added to the 5' end of the primer to include restriction endonuclease cleavage sites to allow for convenient cloning of the PCR fragment into cloning vectors. Addition of such 5' "tails" is well known in the art, and do not interfere with amplification as mismatched nucleotides at the 5' end of primers are not required for amplification.

For example, as demonstrated in the examples that follow, amino acid sequence alignments between sequences encoding 13hydroperoxy lyase from bell pepper (Matsui, et al. (1996) FEBS Letters 394:21-24), banana (European Patent Application, Publication Number EP 0 801 133 A2, the entirety of which is incorporated herein by reference) and Arabidopsis identify highly conserved peptide sequences. Synthetic oligonucleotide primers are constructed based on these peptide sequences and used in PCR reactions to amplify cDNAs from several plant tissue sources which are highly similar to known HPO lyase sequences. Complimentary DNA sequences are identified from tomato hypocotyl and fruit tissues and cucumber hypocotyl tissue which are highly homologous to the bell pepper 13-HPO lyase. However, a second cDNA sequence is obtained from the cucumber hypocotyl tissue which is divergent from the 13-HPO lyase sequences. Furthermore, the sequence shows a slightly higher similarity to allene oxide synthases.

The nucleotide sequence obtained from any method may be molecularly cloned into an appropriate vector for propagation of the DNA by methods known in the art. Many cloning vectors

10

15

20

30

are available commercially, and may be employed used with the nucleotide sequences of the present invention. For example, the DNA may be inserted into a pBluescript vector (Stratgene, La Jolla, CA). Typically, useful cloning vectors for bacterial use can comprise a called 11

- bacterial use can comprise a selectable marker and a bacterial origin of replication derived from commercially available plasmids which are in turn derived from the well known cloning vector pBR322 (ATCC 37017). Further examples of cloning vectors include the pGEM vectors (Promega Biotec, Madison,
- 10 WI). In addition, linear cloning vectors with single thymine ("T") overhangs are available for convenient cloning of PCR fragments amplified utilizing Taq DNA polymerase which usually adds an adenine ("A") to the end of the amplified fragment. Such vectors include pCR2.1 vector (Invitrogen, La Jolla, CA).
- Cloned DNA sequences may be expressed in hosts to confirm HPO lyase activity. Techniques and expression vectors for expression in such hosts are well known in the art and are available from a wide variety of commercial sources. Examples of such vectors include pQE30 (Qiagen, Hilden, Germany).
- Furthermore, the sequences of the DNA coding regions can be optimized by gene resynthesis, based on codon usage, for maximum expression in particular hosts.

As described in more detail in the examples that follow, expression of the cucumber HPO lyase sequence in *E coli* demonstrates that the sequence encodes a 9-hydroperoxide lyase. This is the first reported sequence cloned, to the inventor's knowledge, encoding a hydroperoxide lyase with activity on linolenic acid 9-hydroperoxide substrates.

Once the nucleic acid sequence is obtained, the

transcription, or transcription and translation (expression),
of the 9-HPO lyase in a host cell is desired to produce a
ready source of the enzyme and/or modify the composition of
fatty acids and/or volatile compound found therein. Other
useful applications may be found when the host cell is a plant
host cell, in vitro and in vivo.

Nucleic acids (genomic DNA, plasmid DNA, cDNA, synthetic DNA, mRNA, etc.) encoding HPO lyase or amino acid sequences of the purified enzymes, which permit design of nucleic acid

25

probes facilitating the isolation of DNA coding sequences therefor, are known in the art and are available for use in the methods of the present invention. It is generally recognized to an artisan skilled in the field to which the present invention pertains that the nucleic acid sequences provided herein and the amino acid sequences derived therefrom may be used to isolate other potential HPO lyase genes from GenBank or other database using DNA and peptide search techniques generally known in the art.

In addition to the sequences described in the present invention, DNA coding sequences useful in the present invention can be derived from algae, fungi, bacteria, mammalian sources, plants, etc. Homology searches in existing databases using signature sequences corresponding to conserved nucleotide and amino acid sequences of 9-HPO lyase can be employed to isolate equivalent, related genes from other sources such as plants and microorganisms. Searches in EST Furthermore, the use of DNA databases can also be employed. sequences encoding enzymes functionally enzymatically equivalent to those disclosed herein, wherein such DNA sequences are degenerate equivalents of the nucleic acid sequences disclosed herein in accordance with the degeneracy of the genetic code, is also encompassed by the present invention. Demonstration of the functionality of coding sequences identified by any of these methods can be carried out by expression of such sequences in various expression systems employing various organisms.

Furthermore, constructs for the expression of the 9-HPO lyase sequences in plants are described. Also, the use of such 9-HPO lyase expression constructs for increasing aldehyde and oxo-acid production in plant tissues is contemplated in the present invention. The 9-HPO lyase expression constructs may be employed to produce a host plant having modified organoleptic properties. By organoleptic properties is meant as relating to qualities, such as taste, color, odor, and feel, of a substance, such as a food, that stimulate the sense organs. Furthermore, constructs may be prepared to direct the expression of the 9-HPO lyase sequence in a host cell to

5

10

15

20

25

30

provide for the production of aldehydes and oxo-acids for use in flavorings and fragrances.

The nucleic acid sequences which encode 9-hydroperoxide lyase may be used in various constructs, for example, as probes to obtain further sequences. Alternatively, these sequences may be used in conjunction with appropriate regulatory sequences to increase levels of the respective 9-HPO lyase of interest in a host cell for recovery or study of the enzyme *in vitro* or *in vivo* or to decrease levels of the respective 9-HPO lyase of interest for some applications when the host cell is a plant entity, including plant cells, plant parts (including but not limited to seeds, cuttings or tissues) and plants.

Thus, depending upon the intended use, the constructs may contain the nucleic acid sequence which encodes the entire 9-HPO lyase protein, or a portion thereof. For example, where antisense inhibition of a given 9-HPO lyase protein is desired, the entire 9-HPO lyase sequence is not required. Furthermore, where 9-HPO lyase constructs are intended for use as probes, it may be advantageous to prepare constructs containing only a particular portion of a 9-HPO lyase encoding sequence, for example a sequence which is discovered to encode a highly conserved 9-HPO lyase region.

As discussed above, nucleic acid sequence encoding a

plant or other 9-HPO lyase of this invention may include
genomic, cDNA or mRNA sequence. By "encoding" is meant that
the sequence corresponds to a particular amino acid sequence
either in a sense or anti-sense orientation. By
"extrachromosomal" is meant that the sequence is outside of
the plant genome of which it is naturally associated. By
"recombinant" is meant that the sequence contains a
genetically engineered modification through manipulation via
mutagenesis, restriction enzymes, and the like.

A cDNA sequence may or may not contain pre-processing

sequences, such as transit peptide sequences or targeting
sequences to facilitate delivery of the 9-HPO lyase protein
to a given organelle or membrane location. The use of any
such precursor 9-HPO lyase DNA sequences is preferred for uses

10

15

in plant cell expression. A genomic 9-HPO lyase sequence may contain the transcription and translation initiation regions, introns, and/or transcript termination regions of the plant 9-HPO lyase, which sequences may be used in a variety of DNA constructs, with or without the 9-HPO lyase structural gene. Thus, nucleic acid sequences corresponding to the 9-HPO lyase of this invention may also provide signal sequences useful to direct protein delivery into a particular organellar or membrane location, 5' upstream non-coding regulatory regions (promoters) having useful tissue and timing profiles, 3' downstream non-coding regulatory regions useful as transcriptional and translational regulatory regions, and may lend insight into other features of the gene.

Once the desired plant or other 9-HPO lyase nucleic acid sequence is obtained, it may be manipulated in a variety of ways. Where the sequence involves non-coding flanking regions, the flanking regions may be subjected to resection, mutagenesis, etc. Thus, transitions, transversions, deletions, and insertions may be performed on the naturally occurring sequence. In addition, all or part of the sequence may be synthesized. In the structural gene, one or more codons may be modified to provide for a modified amino acid sequence, or one or more codon mutations may be introduced to provide for a convenient restriction site or other purpose involved with construction or expression. The structural gene **25** . may be further modified by employing synthetic adapters, linkers to introduce one or more convenient restriction sites, or the like.

The nucleic acid or amino acid sequences encoding a plant or other 9-HPO lyase of this invention may be combined with other non-native, or "heterologous", sequences in a variety of ways. By "heterologous" sequences is meant any sequence which is not naturally found joined to the native (or wild-type) 9-HPO lyase, including, for example, combinations of nucleic acid sequences from the same plant which are not naturally found joined together. A state of the state

The DNA sequence encoding a plant or other 9-HPO lyase of this invention may be employed in conjunction with all or part

. 5

10

15

20

30

WO 00/50575 14 PCT/US00/05311

of the gene sequences normally associated with the 9-HPO lyase. In its component parts, a DNA sequence encoding 9-HPO lyase is combined in a DNA construct having, in the 5' to 3' direction of transcription, a transcription initiation control region capable of promoting transcription and translation in a host cell, the DNA sequence encoding plant 9-HPO lyase and a transcription and translation termination region.

Potential host cells include both prokaryotic and eukaryotic cells. A host cell may be unicellular or found in a multicellar differentiated or undifferentiated organism depending upon the intended use. Cells of this invention may be distinguished by having a 9-HPO lyase foreign to the wild-type cell present therein, for example, by having a recombinant nucleic acid construct encoding a 9-HPO lyase therein not native to the host species.

Prokaryotic cells include gram negative as well as gram positive bacteria, for example E. coli, and B. subtilis strains. Suitable examples are well known to the skilled artisan. As described in more detail in the examples that follow, an HPO lyase isolated from cucumber hypocotyl is expressed in E coli, strain M15. The protein expressed from the E coli is capable of producing the aldehyde 3(Z)-nonenal and 2(E)-nonenal from linoleic acid 9-hydroperoxide. Thus, the HPO lyase isolated from the cucumber hypocotyl encodes a 9-HPO lyase.

Eukaryotic host cells include fungi, including yeasts, insect cells, and plant cells. Methods for the expression of DNA sequences of interest in yeast cells are known in the art and are generally described in "Guide to yeast genetics and molecular biology", Guthrie and Fink, eds. Methods in enzymology, Academic Press, Inc. Vol 194 (1991) and Gene expression technology", Goeddel ed, Methods in Enzymology, Academic Press, Inc., Vol 185 (1991). In addition, methods for the expression of 13-HPO lyase genes are described in European patent Application EP 0 801 133 A2, the entirety of which is incorporated herein by reference.

The fungal recombinant vector may be any vector which can be conveniently subjected to recombinant DNA procedures. The

30

35

10

choice of a vector will typically depend on the compatibility of the vector with the fungal host cell into which the vector is to be introduced. The vector may be a linear or a closed circular plasmid. The vector system may be a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the fungal host.

The fungal vector may be an autonomously replicating vector, i.e., a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced 15 into the fungal cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. For integration, the vector may rely on the nucleic acid sequence of the vector for stable integration of the vector into the genome by homologous or nonhomologous 20 recombination. Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the fungal host. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of 25 integration at a precise location, there should be preferably two nucleic acid sequences which individually contain a sufficient number of nucleic acids, preferably 400bp to 1500bp, more preferably 800bp to 1000bp, which are highly 30 homologous with the corresponding target sequence to enhance the probability of homologous recombination. These nucleic acid sequences may be any sequence that is homologous with a target sequence in the genome of the fungal host cell and, furthermore, may be non-encoding or encoding sequences.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of origin of replications for use in a yeast host cell are the

35

WO 00/50575 16 PCT/US00/05311

2 micron origin of replication and the combination of CEN3 and ARS 1. Any origin of replication may be used which is compatible with the fungal host cell of choice.

The fungal vectors of the present invention preferably contain one or more selectable markers which permit easy 5 selection of transformed cells. A selectable marker is a gene the product of which provides, for example biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs and the like. The selectable marker may be selected from the group including, but not limited to, amdS 10 (acetamidase), argB (ornithine carbamoyltransferase), bar (phosphinothricin acetyltransferase), hygB (hygromycin phosphotransferase), niaD (nitrate reductase), pyrG (orotidine-5'-phosphate decarboxylase) and sC (sulfate 15 adenyltransferase) and trpC (anthranilate synthase). Preferred for use in an Aspergillus cell are the amdS and pyrG markers of Aspergillus nidulans or Aspergillus oryzae and the bar marker of Streptomyces hygroscopicus. Furthermore, selection may be accomplished by co-transformation, e.g., as described in WO 91/17243, the entirety of which is herein 20 incorporated by reference. A nucleic acid sequence of the present invention may be operably linked to a suitable promoter sequence. The promoter sequence is a nucleic acid sequence which is recognized by the fungal host cell for expression of the nucleic acid sequence. 25 The promoter sequence contains transcription and translation control sequences which mediate the expression of the protein or fragment thereof.

transcriptional activity in the fungal host cell of choice and may be obtained from genes encoding polypeptides either homologous or heterologous to the host cell. Examples of suitable promoters for directing the transcription of a nucleic acid construct of the invention in a filamentous fungal host are promoters obtained from the genes encoding Aspergillus oryzae TAKA amylase, Rhizomucor miehei aspartic proteinase, Aspergillus niger neutral alpha-amylase, Aspergillus niger acid stable alpha-amylase, Aspergillus niger

or Aspergillus awamori glucoamylase (glaA), Rhizomucor miehei lipase, Aspergillus oryzae alkaline protease, Aspergillus oryzae triose phosphate isomerase, Aspergillus nidulans acetamidase and hybrids thereof. In a yeast host, a useful promoter is the Saccharomyces cerevisiae enolase (eno-1) promoter. Particularly preferred promoters are the TAKA amylase, NA2-tpi (a hybrid of the promoters from the genes encoding Aspergillus niger neutral alpha -amylase and Aspergillus oryzae triose phosphate isomerase) and glaA promoters.

A protein or fragment thereof encoding nucleic acid molecule of the present invention may also be operably linked to a terminator sequence at its 3' terminus. The terminator sequence may be native to the nucleic acid sequence encoding the protein or fragment thereof or may be obtained from foreign sources. Any terminator which is functional in the fungal host cell of choice may be used in the present invention, but particularly preferred terminators are obtained from the genes encoding Aspergillus oryzae TAKA amylase, Aspergillus niger glucoamylase, Aspergillus nidulans anthranilate synthase, Aspergillus niger alpha-glucosidase and Saccharomyces cerevisiae enolase.

A protein or fragment thereof encoding nucleic acid molecule of the present invention may also be operably linked to a suitable leader sequence. A leader sequence is a nontranslated region of a mRNA which is important for translation by the fungal host. The leader sequence is operably linked to the 5' terminus of the nucleic acid sequence encoding the protein or fragment thereof. The leader sequence may be native to the nucleic acid sequence encoding the protein or fragment thereof or may be obtained from foreign sources. Any leader sequence which is functional in the fungal host cell of choice may be used in the present invention, but particularly preferred leaders are obtained from the genes encoding Aspergillus oryzae TAKA amylase and Aspergillus oryzae triose phosphate isomerase.

The 3' terminus of the nucleic acid sequence of the present

5

10

15

20

25

30

WO 00/50575 18 PCT/US00/05311

invention. The polyadenylation sequence is a sequence which when transcribed is recognized by the fungal host to add polyadenosine residues to transcribed mRNA. The polyadenylation sequence may be native to the nucleic acid sequence encoding the protein or fragment thereof or may be obtained from foreign sources. Any polyadenylation sequence which is functional in the fungal host of choice may be used in the present invention, but particularly preferred polyadenylation sequences are obtained from the genes encoding Aspergillus oryzae TAKA amylase, Aspergillus niger glucoamylase, Aspergillus nidulans anthranilate synthase and Aspergillus niger alpha-glucosidase.

To avoid the necessity of disrupting the cell to obtain the protein or fragment thereof and to minimize the amount of possible degradation of the expressed protein or fragment 15 thereof within the cell, it is preferred that expression of the protein or fragment thereof gives rise to a product secreted outside the cell. To this end, a protein or fragment thereof of the present invention may be linked to a signal peptide linked to the amino terminus of the protein or 20 fragment thereof. A signal peptide is an amino acid sequence which permits the secretion of the protein or fragment thereof from the fungal host into the culture medium. The signal peptide may be native to the protein or fragment thereof of the invention or may be obtained from foreign sources. 25 end of the coding sequence of the nucleic acid sequence of the present invention may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted protein or fragment thereof. Alternatively, the 5' 30 end of the coding sequence may contain a signal peptide coding region which is foreign to that portion of the coding sequence which encodes the secreted protein or fragment thereof. foreign signal peptide may be required where the coding sequence does not normally contain a signal peptide coding 35 Alternatively, the foreign signal peptide may simply replace the natural signal peptide to obtain enhanced secretion of the desired protein or fragment thereof.

foreign signal peptide coding region may be obtained from a glucoamylase or an amylase gene from an Aspergillus species, a lipase or proteinase gene from Rhizomucor miehei, the gene for the alpha-factor from Saccharomyces cerevisiae, or the calf preprochymosin gene. An effective signal peptide for fungal host cells is the Aspergillus oryzae TAKA amylase signal, Aspergillus niger neutral amylase signal, the Rhizomucor miehei aspartic proteinase signal, the Humicola lanuginosus cellulase signal, or the Rhizomucor miehei lipase signal. However, any signal peptide capable of permitting secretion of the protein or fragment thereof in a fungal host of choice may be used in the present invention.

A protein or fragment thereof encoding nucleic acid molecule of the present invention may also be linked to a propeptide coding region. A propeptide is an amino acid sequence found at the amino terminus of aproprotein or proenzyme. Cleavage of the propeptide from the proprotein yields a mature biochemically active protein. The resulting polypeptide is known as a propolypeptide or proenzyme (or a zymogen in some cases). Propolypeptides are generally inactive and can be converted to mature active polypeptides by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide or proenzyme. The propeptide coding region may be native to the protein or fragment thereof or may be obtained from foreign sources. The foreign propeptide coding region may be obtained from the Saccharomyces cerevisiae alpha-factor gene or Myceliophthora thermophila laccase gene (WO 95/33836, the entirety of which is herein incorporated by reference).

The procedures used to ligate the elements described above to construct the recombinant expression vector of the present invention are well known to one skilled in the art (see, for example, Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd ed., Cold Spring Harbor, N.Y., (1989)).

The present invention also relates to recombinant fungal host cells produced by the methods of the present invention which are advantageously used with the recombinant vector of the present invention. The cell is preferably transformed

5

10

15

20

30

with a vector comprising a nucleic acid sequence of the invention followed by integration of the vector into the host chromosome. The choice of fungal host cells will to a large extent depend upon the gene encoding the protein or fragment thereof and its source. The fungal host cell may, for example, be a yeast cell or a filamentous fungal cell.

"Yeast" as used herein includes Ascosporogenous yeast (Endomycetales), Basidiosporogenous yeast and yeast belonging to the Fungi Imperfecti (Blastomycetes). The Ascosporogenous yeasts are divided into the families Spermophthoraceae and 10 Saccharomycetaceae. The latter is comprised of four subfamilies, Schizosaccharomycoideae (for example, genus Schizosaccharomyces), Nadsonioideae, Lipomycoideae and Saccharomycoideae (for example, genera Pichia, Kluyveromyces and Saccharomyces). The Basidiosporogenous yeasts include the 15 genera Leucosporidim, Rhodosporidium, Sporidiobolus, Filobasidium and Filobasidiella. Yeast belonging to the Fungi Imperfecti are divided into two families, Sporobolomycetaceae (for example, genera Sorobolomyces and Bullera) and 20 Cryptococcaceae (for example, genus Candida). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in Biology and Activities of Yeast (Skinner et al., Soc. App. Bacteriol. Symposium Series No. 9, (1980), the entirety of which is herein incorporated by reference). 25 biology of yeast and manipulation of yeast genetics are well known in the art (see, for example, Biochemistry and Genetics of Yeast, Bacil et al. (ed.), 2nd edition, 1987; The Yeasts, Rose and Harrison (eds.), 2nd ed., (1987); and The Molecular 30 Biology of the Yeast Saccharomyces, Strathern et al. (eds.), (1981), all of which are herein incorporated by reference in their entirety).

The recombinant fungal host cells of the present invention may further comprise one or more sequences which encode one or more factors that are advantageous in the expression of the protein or fragment thereof, for example, an activator (e.g., a trans-acting factor), a chaperone and a processing protease. The nucleic acids encoding one or more

35

of these factors are preferably not operably linked to the nucleic acid encoding the protein or fragment thereof. activator is a protein which activates transcription of a nucleic acid sequence encoding a polypeptide (Kudla et al., 5 EMBO 9:1355-1364(1990); Jarai and Buxton, Current Genetics 26:2238-244(1994); Verdier, Yeast 6:271-297(1990), all of which are herein incorporated by reference in their entirety). The nucleic acid sequence encoding an activator may be obtained from the genes encoding Saccharomyces cerevisiae heme 10 activator protein 1 (hap1), Saccharomyces cerevisiae galactose metabolizing protein 4 (gal4) and Aspergillus nidulans ammonia regulation protein (areA). For further examples, see Verdier, Yeast 6:271-297 (1990); MacKenzie et al., Journal of Gen. Microbiol. 139:2295-2307 (1993), both of which are herein 15 incorporated by reference in their entirety). A chaperone is a protein which assists another protein in folding properly (Hartl et al., TIBS 19:20-25 (1994); Bergeron et al., TIBS 19:124-128 (1994); Demolder et al., J. Biotechnology 32:179-189 (1994); Craig, Science 260:1902-1903(1993); Gething and 20 Sambrook, Nature 355:33-45 (1992); Puig and Gilbert, J Biol. Chem. 269:7764-7771 (1994); Wang and Tsou, FASEB Journal 7:1515-11157 (1993); Robinson et al., Bio/Technology 1:381-384 (1994), all of which are herein incorporated by reference in their entirety). The nucleic acid sequence encoding a chaperone may be obtained from the genes encoding Aspergillus oryzae protein disulphide isomerase, Saccharomyces cerevisiae calnexin, Saccharomyces cerevisiae BiP/GRP78 and Saccharomyces cerevisiae Hsp70. For further examples, see Gething and Sambrook, Nature 355:33-45 (1992); Hartl et al., TIBS 19:20-25 30 (1994). A processing protease is a protease that cleaves a propeptide to generate a mature biochemically active polypeptide (Enderlin and Ogrydziak, Yeast 10:67-79 (1994); Fuller et al., Proc. Natl. Acad. Sci. (U.S.A.) 86:1434-1438 (1989); Julius et al., Cell 37:1075-1089 (1984); Julius et 3.5 al., Cell 32:839-852 (1983), all of which are incorporated by reference in their entirety). The nucleic acid sequence encoding a processing protease may be obtained from the genes encoding Aspergillus niger Kex2, Saccharomyces cerevisiae

WO 00/50575 22 PCT/US00/05311

dipeptidylaminopeptidase, Saccharomyces cerevisiae Kex2 and Yarrowia lipolytica dibasic processing endoprotease (xpr6). Any factor that is functional in the fungal host cell of choice may be used in the present invention.

Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts and regeneration of the cell wall in a manner known per se. Suitable procedures for transformation of Aspergillus host cells are described in EP 238 023 and Yelton et al., Proc.

Natl. Acad. Sci. (U.S.A.) 81:1470-1474 (1984), both of which are herein incorporated by reference in their entirety. A suitable method of transforming Fusarium species is described by Malardier et al., Gene 78:147-156 (1989), the entirety of which is herein incorporated by reference. Yeast may be

transformed using the procedures described by Becker and Guarente, In: Abelson and Simon, (eds.), Guide to Yeast Genetics and Molecular Biology, Methods Enzymol. Volume 194, pp 182-187, Academic Press, Inc., New York; Ito et al., J. Bacteriology 153:163 (1983); Hinnen et al., Proc. Natl. Acad.

20 Sci. (U.S.A.) 75:1920 (1978), all of which are herein incorporated by reference in their entirety.

The present invention also relates to methods of producing the protein or fragment thereof comprising culturing the recombinant fungal host cells under conditions conducive for expression of the protein or fragment thereof. 25 The fungal cells of the present invention are cultivated in a nutrient medium suitable for production of the protein or fragment thereof using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-30 batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the protein or fragment thereof to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen 35 sources and inorganic salts, using procedures known in the art (see, e.g., Bennett and LaSure (eds.), More Gene Manipulations in Fungi, Academic Press, CA, (1991), the entirety of which is

herein incorporated by reference). Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection, Manassas, VA). If the protein or fragment thereof is secreted into the nutrient medium, a protein or fragment thereof can be recovered directly from the medium. If the protein or fragment thereof is not secreted, it is recovered from cell lysates.

The expressed protein or fragment thereof may be detected using methods known in the art that are specific for the particular protein or fragment. These detection methods may include the use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, if the protein or fragment thereof has enzymatic activity, an enzyme assay may be used. Alternatively, if polyclonal or monoclonal antibodies specific to the protein or fragment thereof are available, immunoassays may be employed using the antibodies to the protein or fragment thereof. The techniques of enzyme assay and immunoassay are well known to those skilled in the art.

The resulting protein or fragment thereof may be recovered by methods known in the arts. For example, the protein or fragment thereof may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation. The recovered protein or fragment thereof may then be further purified by a variety of chromatographic procedures, e.g., ion exchange chromatography, gel filtration chromatography, affinity chromatography, or the like.

Methods for the expression of DNA sequences of interest in insect host cells are also well known in the art, and are reviewed by Lucow and Summers, (1988) *Bio/technology* 6:47-55, the entirety of which is incorporated herein by reference.

In a preferred embodiment, plant host cells are employed in transformation experiments with constructs containing DNA sequences coding 9-HPO lyase enzymes expressed from constitutive or tissue enhanced promoters.

5

10

15

20

25

30

WO 00/50575 24 PCT/US00/05311

Depending upon the host, the regulatory regions will vary, including regions from viral, plasmid or chromosomal genes, or the like. For expression in prokaryotic or eukaryotic microorganisms, particularly unicellular hosts, a wide variety of constitutive or regulatable promoters may be employed. Expression in a microorganism can provide a ready source of the plant enzyme. Among transcriptional initiation regions which have been described are regions from bacterial and yeast hosts, such as *E. coli*, *B. subtilis*, *Sacchromyces cerevisiae*, including genes such as beta-galactosidase, T7 polymerase, tryptophan E and the like.

In a preferred embodiment, the constructs will involve regulatory regions functional in plants which provide for modified production of 9-HPO lyase, and, possibly,

frame coding for the plant HPO lyase or functional fragment thereof will be joined at its 5' end to a transcription initiation regulatory region. In embodiments wherein the expression of the 9-HPO lyase protein is desired in a plant 20 host, the use of 211 to make a finite composition.

host, the use of all or part of the complete plant HPO lyase gene is desired; namely all or part of the 5' upstream non-coding regions (promoter) together with the structural gene sequence and 3' downstream non-coding regions may be employed.

If a different promoter is desired, such as a promoter 25 native to the plant host of interest or a modified promoter, i.e., having transcription initiation regions derived from one gene source and translation initiation regions derived from a different gene source, numerous transcription initiation regions are available which provide for a wide variety of 30 constitutive or regulatable, e.g., inducible, transcription of the structural gene functions. The transcription/translation initiation regions corresponding to such structural genes are found immediately 5' upstream to the respective start codons. Among transcriptional initiation regions used for plants are such regions associated with the T-DNA structural genes such 35 as for nopaline and mannopine synthases, the 19S and 35S promoters from CaMV, and the 5' upstream regions from other

plant genes such as napin, ACP, SSU, PG, zein, phaseolin E,

and the like. Enhanced promoters, such as double 35S, are also available for expression of HPO lyase sequences. For such applications when 5' upstream non-coding regions are obtained from other genes regulated during seed maturation, those preferentially expressed in plant embryo tissue, such as ACP and napin-derived transcription initiation control regions, are desired. Such "seed-specific promoters" may be obtained and used in accordance with the teachings of issued U.S. Patent Numbers 5,608,152 and 5,530,194, which references are hereby incorporated by reference. Transcription initiation regions which are preferentially expressed in seed tissue, i.e., which are undetectable in other plant parts, are considered desirable for expression in order to minimize any disruptive or adverse effects of the gene product.

Regulatory transcript termination regions may be provided in DNA constructs of this invention as well. Transcript termination regions may be provided by the DNA sequence encoding the plant HPO lyase or a convenient transcription termination region derived from a different gene source, for example, the transcript termination region which is naturally associated with the transcript initiation region. Where the transcript termination region is from a different gene source, it will contain at least about 0.25 kb, preferably about 1-3 kb of sequence 3' to the structural gene from which the termination region is derived.

Plant expression or transcription constructs having a 9-HPO lyase as the DNA sequence of interest for increased or decreased expression thereof may be employed with a wide variety of plant life, particularly, plant life involved in the production of tissues for edible uses. Most especially preferred are seed, fruit, vegetable and leaf crops. Plants of interest include, but are not limited to, Brassica species, soybean, corn, tomato, strawberry, bell pepper and melon. Depending on the method for introducing the recombinant constructs into the host cell, other DNA sequences may be required. Importantly, this invention is applicable to dicotyledenous and monocotyledenous species alike and will be

5

10

15

20

25

30

WO 00/50575 26 PCT/US00/05311

readily applicable to new and/or improved transformation and regulation techniques.

In order to increase lipid peroxidation, and thereby increasing "melon" flavors/fragrances, in a plant tissue,

5 coexpression of a plant or other 9-HPO lyase in a plant tissue with a second gene involved in lipid peroxidation may also find use in the present invention. For example, coexpression of a 9-HPO lyase sequence in a plant tissue with a DNA sequence encoding for another protein involved in lipid peroxidation, such as a lipoxygenase may increase lipid peroxidation and increase the total short-chain aldehydes produced in the plant tissue. Such an increase in short-chain aldehydes may increase the "melon" flavor in an edible plant tissue.

15 The method of transformation in obtaining such transgenic plants is not critical to the instant invention, and various methods of plant transformation are currently available. Furthermore, as newer methods become available to transform crops, they may also be directly applied hereunder. example, many plant species naturally susceptible to 20 Agrobacterium infection may be successfully transformed via tripartite or binary vector methods of Agrobacterium mediated transformation. In many instances, it will be desirable to have the construct bordered on one or both sides by T-DNA, particularly having the left and right borders, more 25 particularly the right border. This is particularly useful when the construct uses A. tumefaciens or A. rhizogenes as a mode for transformation, although the T-DNA borders may find use with other modes of transformation. In addition, techniques of microinjection, DNA particle bombardment, and 30 electroporation have been developed which allow for the transformation of various monocot and dicot plant species.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformant cells. The gene may provide for resistance to a cytotoxic agent, e.g. antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an auxotrophic host,

viral immunity or the like. Depending upon the number of different host species the expression construct or components thereof are introduced, one or more markers may be employed, where different conditions for selection are used for the different hosts.

Where Agrobacterium is used for plant cell transformation, a vector may be used which may be introduced into the Agrobacterium host for homologous recombination with T-DNA or the Ti- or Ri-plasmid present in the Agrobacterium host. The Ti- or Ri-plasmid containing the T-DNA for recombination may be armed (capable of causing gall formation) or disarmed (incapable of causing gall formation), the latter being permissible, so long as the vir genes are present in the transformed Agrobacterium host. The armed plasmid can give a mixture of normal plant cells and gall.

In some instances where Agrobacterium is used as the vehicle for transforming host plant cells, the expression or transcription construct bordered by the T-DNA border region(s) will be inserted into a broad host range vector capable of replication in E. coli and Agrobacterium, there being broad host range vectors described in the literature. Commonly used is pRK2 or derivatives thereof. See, for example, Ditta, et al., (Proc. Nat. Acad. Sci., U.S.A. (1980) 77:7347-7351) and EPA 0 120 515, which are incorporated herein by reference. Alternatively, one may insert the sequences to be expressed in plant cells into a vector containing separate replication sequences, one of which stabilizes the vector in E. coli, and the other in Agrobacterium. See, for example, McBride and Summerfelt (Plant Mol. Biol. (1990) 14:269-276), wherein the pRiHRI (Jouanin, et al., Mol. Gen. Genet. (1985) 201:370-374) origin of replication is utilized and provides for added stability of the plant expression vectors in host Agrobacterium cells.

Included with the expression construct and the T-DNA will
be one or more markers, which allow for selection of
transformed Agrobacterium and transformed plant cells. A
number of markers have been developed for use with plant
cells, such as resistance to chloramphenicol, kanamycin, the

5

10

15

20

WO 00/50575 28 PCT/US00/05311

aminoglycoside G418, hygromycin, or the like. The particular marker employed is not essential to this invention, one or another marker being preferred depending on the particular host and the manner of construction.

For transformation of plant cells using Agrobacterium, explants may be combined and incubated with the transformed Agrobacterium for sufficient time for transformation, the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once callus forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the shoots transferred to rooting medium for regeneration of plants. The plants may then be grown to seed and the seed used to establish repetitive generations and for isolation of vegetable oils.

Once transgenic hosts have been obtained which express the 9-HPO lyase encoding sequence, a number of methods are available in the art for the analysis of the expression and protein activity. Methods for the analysis of 9-HPO lyase expression, such as Southern and Northern hybridizations and Western immuno blot assays are generally described in Maniatis, et al. (1989) Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. Methods for the analysis of HPO lyase activity are described in Matsui, et al. (1996) FEBS Letters 394:21-24; Rehbock, et al. (1998) Fresenius J. Analytical Chem 360:830-832; Nunez, et al. (1998) Lipids 33(5):533-538; and in European Patent Application EP 0 801 133

Host plants expressing the 9-HPO lyase sequences of the present invention find use in the production of volatile aldehydes referred to as melon note compounds. Such compounds can be used to sharpen and enhance flavored products, such as those having fruit flavors. Such compounds and uses are described by Morris, (1981) Perfumer & Flavorist vol 6 nol, and Clark, (1990) Perfumer and Flavorist vol 15, the entireties of which are incorporated herein by reference.

10

Host cells expressing the 9-HPO lyases of the present invention provide a novel source of volatile aldehydes contributing to "melon" notes for use in various applications. Furthermore, the host cells may also contain constructs providing for a increased production of enzymes involved in lipid peroxidation, for example lipoxygenase. In addition, the host cells may also produce an increased amount of a particular fatty acid, or have a general increase in fatty acids. Such host cells may be obtained using traditional breeding techniques, including mutagenesis, as well as hosts genetically engineered with such an altered fatty acid composition.

Furthermore, plant host cells containing a construct providing for the expression of the 9-HPO lyase sequences of the present invention find use as a source for aldehydes in reactions for the production of alcohols for use in flavorings and aromatic products. Such methods are known in the art and are described for example in U.S. Patent Number 5,695,973 and in PCT Publication WO 95/26413 the entireties of which are incorporated herein by reference. Generally, a mixture of aldehydes and alcohols are obtained from such methods. The methods generally involve a reaction mixture containing at least one unsaturated fatty acid, a plant material having a relatively high amount of enzyme activity of lipoxygenase and hydroperoxide lyase, and a source of alcohol dehydrogenase.

The unsaturated fatty acid may vary and include a single unsaturated fatty acid species as well as mixtures of several unsaturated fatty acids. The fatty acids are provided in a free acid form, and examples include, but are not limited to oleic acid, linoleic acid, linolenic acid (alpha and gamma forms), arachidonic acid, eicosapentaenoic acid, and ricinoleic acid.

Sources of the alcohol dehydrogenase include yeasts, as well as non-yeast molds. The alcohol dehydrogenase has the ability to convert an aldehyde to an alcohol. The yeast and non-yeast molds further provide a source of nicotine adenine dinucleotide (NADH) as a reducing agent.

10

15

20

25

30

WO 00/50575 30 PCT/US00/05311

The invention now being generally described, it will be more readily understood by reference to the following examples which are included for purposes of illustration only and are not intended to limit the present invention.

5

10

15

20

EXAMPLES

Example 1 Identification of Arabidopsis HPO Lyase Sequences
A nucleic acid fragment encoding hydroperoxide lyase from
Bell pepper has been previously cloned and sequenced (Matsui,
et al. (1996) supra). The nucleotide sequence was used to
search Genbank for HPO lyase related sequences. One accession
identified from Genbank (Accession Number Z97339,
(http://www.ncbi.nlm.nih.gov/web/Genbank/Index.html))
containing a genomic sequence from Arabidopsis was reported to
encode an allene oxide synthase.

Sequence comparisons between the bell pepper HPO lyase, Arabidopsis allene oxide synthase (Laudert, et al. (1996) supra) and the Arabidopsis HPO lyase-like sequence from Genbank using Genetyx Mac (Software Development Co. Ltd.) indicated that the Arabidopsis HPO lyase-like sequence is more similar to the bell pepper HPO lyase (57% identity) (see Figure 2) than to the allene oxide synthase sequence (39% identity) (see Figure 3).

Albander Carry Carry Example 2 Construction of Arabidopsis cDNA libraries 25 Total RNA from seedling, inflorescence, and silique tissues of Arabidopsis thaliana is isolated for use in construction of complementary (cDNA) libraries. The procedure is an adaptation of the DNA isolation protocol of Webb and 30 Knapp (D.M. Webb and S.J. Knapp, (1990) Plant Molec. Reporter, 8, 180-185). The following description assumes the use of 1g fresh weight of tissue. Frozen seed tissue is powdered by grinding under liquid nitrogen. The powder is added to 10ml REC buffer (50mM Tris-HCl, pH 9, 0.8M NaCl, 10mM EDTA, 0.5% w/v CTAB (cetyltrimethyl-ammonium bromide)) along with 0.2g 35 insoluble polyvinylpolypyrrolidone, and ground at room temperature. The homogenate is centrifuged for 5 minutes at $12,000 ext{ x}g$ to pellet insoluble material. The resulting

supernatant fraction is extracted with chloroform, and the top phase is recovered.

The RNA is then precipitated by addition of 1 volume RecP (50mM Tris-HCL pH9, 10mM EDTA and 0.5% (w/v) CTAB) and collected by brief centrifugation as before. The RNA pellet is redissolved in 0.4 ml of 1M NaCl. The RNA pellet is redissolved in water and extracted with phenol/chloroform. Sufficient 3M potassium acetate (pH 5) is added to make the mixture 0.3M in acetate, followed by addition of two volumes of ethanol to precipitate the RNA. After washing with ethanol, this final RNA precipitate is dissolved in water and stored frozen.

Alternatively, total RNAs may be obtained using TRIzol reagent (BRL Life Technologies, Gaithersburg, MD) following the manufacturers protocol.

Complementary DNAs (cDNA) are obtained from the RNAs using the Marathon cDNA Amplification Kit (Clontech, Palo Alto, CA) following the manufacturers directions.

20 Example 3 Cloning of HPO Lyase Sequences

In order to characterize the protein encoded by the Arabidopsis cDNA GenBank sequence, the entire coding region corresponding to the Arabidopsis HPO lyase-like cDNA was obtained. (Figure 1) Synthetic oligo-nucleotide primers are designed to amplify the 5' and 3' ends from the HPO lyase-like sequence from RNA obtained in Example 2. Primers are designed according to the Arabidopsis HPO lyase-like sequence and are used in Rapid Amplification of cDNA Ends (RACE) reactions (Frohman et al. (1988) Proc. Natl. Acad. Sci. USA 85:8998-9002). Amplification of flanking sequences from cDNA clones are performed using the Marathon cDNA Amplification kit (Clontech) according to the manufacturers protocol.

A pair of primers were designed to amplify the 5' and 3' regions from the Arabidopsis HPO lyase-like cDNA from the libraries described in example 2 above. These two primers, HPOL28 (for 3' RACE, 5'-CGGTTCCTCTGCGCCTCTCTCGCCGGCG-3') and HPOL21 (for 5' RACE, 5'-GCGGAACCGGAGGACTAAAACGCAGC-3') are used in PCR reactions with Adapter specific primers (AP1 5'-

5

10

15

25

30

CCATCCTAATACGACTCACTATAGGGC-3') provided in the Marathon cDNA Amplification Kit. For amplification of the 5' region of the HPO lyase-like cDNA the primers AP1 and HPOL 21 were used, and for the amplification of the 3' region the AP1 primer was used in a reaction with the primer HPOL28. The cycle conditions used are: 94°C for 1 minute followed by 5 rounds of 94°C for 5 seconds, 72°C for 4 minutes, followed by 5 rounds of amplification using 94°C for 5 seconds, 70°C for 4 minutes, and finally 25 cycles of 94°C for 5 seconds, and 68°C for 4 minutes.

A single fragment of 1100 bp was obtained from the 3'RACE reaction with RNA obtained from the silique tissue described above. To confirm that the PCR product contained sequence corresponding to the HPO lyase-like sequence, a second round of PCR reactions using the same conditions described above was performed with the gel purified 1100 bp fragment. A reaction was performed with the primers HPOL13 (5'-CTTGGCGTAGTTCCTCAGCCTCTTG-3') and AP2 (5'-ACTCACTATAGGGCTCGAGCGGC-3') to amplify an approximately 1000 bp fragment as a confirmation of the HPO lyase-like sequence. The reamplified 1000 bp fragment was gel purified and cloned into the pCR2.1 TOPO vector (Invitrogen, Carlsbad, CA) to create the plasmid pCGN8094.

The 5' RACE reaction produced many non-specific

fragments. A 1000 bp fragment was excised from the gel and cloned into the pCR2.1 TOPO(Invitrogen) cloning vector to create the plasmid pCGN8091.

Figure 1 discloses the complete nucleotide sequence of the Arabidopsis HPO lyase.

Example 4 Preparation HPO Lyase Expression Constructs

A set of constructs are prepared for transformation into either plant or bacterial hosts to further characterize the Arabidopsis HPO lyase-like sequence. The 5' RACE product in pCGN8091 was PCR amplified using the primers Alex2 (5'-CGGGATCCATGTTGTTGAGAACGATGGCGGCG-3') and Alex4 (5'-CAATCTCCGGCGTTCTCGTCG-3'). The Alex2 primer contains the

5

10

15

20

30

restriction endonuclease site BamHI for the convenient cloning of the PCR product into the pQE30 expression vector (Qiagen, Hilden, Germany) in frame with the ATG start codon of the vector. In addition to the oligonucleotide primers (0.2μM each), the PCR reaction mix contained 0.2 mM each of dATP, dCTP, dGTP and dTTP, 1.0% glycerol, 0.2 mM Tris-HCl (pH 8.3), 4.6 mM KCl, 1.5 mM EDTA, 15 µM dithiothreitol, 7.3 µgm/ml BSA, 1.1 mM KOAc and 0.1 units Pfu DNA polymerase (BRL Life Technologies, Gaithersburg, MD). The mixtures were amplified using the following conditions: 1 cycle of 95°C for 10 minutes; 30 cycles of 94°C for 20 seconds, 60°C for 30 seconds, and 72°C for 1.5 minutes; and, 1 cycle of 72°C for 7 minutes in a Perkin-Elmer 9800 thermocycler. The resulting PCR product was digested with BamHI and HindIII and ligated into the vector pQE30 to create the vector pCGN8099. The 3' terminus of the Arabidopsis HPO lyase was cloned into the HindIII site of pCGN8099 from pCGN8094 to create the E. coli expression vector pCGN8100

A binary vector for plant transformation, pCGN5138, was constructed from pCGN1558 (McBride and Summerfelt, (1990) Plant Molecular Biology, 14:269-276). The polylinker of pCGN1558 was replaced as an *HindIII/EcoRI* fragment with a polylinker containing unique restriction endonuclease sites, *HindIII*, *SseI/PstI*, *NotI*, *BamHI*, *SwaI*, *XbaI*, *PacI*, *AscI*, and *Asp718*.

An antisense construct of the Arabidopsis HPO lyase-like nucleotide sequence was prepared for transformation of Arabidopsis. The nucleic acid sequence encoding the 5' 1000 bp nucleotides from pCGN8091 were cloned as an EcoRI fragment into the plasmid pBluescript II SK (Stratagene, La Jolla, CA) to create the vector pCGN8093. The 3' RACE product from pCGN8090 was cloned as a HindIII fragment into pCGN8093 to create a full length HPO lyase coding sequence in the plasmid pCGN8094. The KpnI site of pCGN8094 was removed by digesting with KpnI and filling in the site with Klenow fragment, and the HPO lyase coding sequence was cloned from this plasmid as a SmaI fragment into the StuI site of pCGN8059. This yields

5

10

15

20

25

30

the plasmid pCGN8101. The plasmid pCGN8059 contains a multiple cloning site downstream of the 35S promoter and the hsp70 leader sequence to allow for the cloning of sequences for expression from the 35S promoter sequence. This vector also contains the nopaline synthase transcription termination (nos 3') sequences (Fraley et al., Proc. Natl. Acad. Sci (1983) 80:4803-4807 and Depicker et al., J. Molec. Appl. Genet. (1982) 1: 562-573). The fragment containing the 35S promoter/hsp70 leader, antisense Arabidopsis HPO lyase sequence, and nos3' termination sequence was cloned from pCGN8101 as a NotI fragment into the same site of pCGN5138 to create the antisense expression construct pCGN8102.

Example 5 E coli Expression

The expression vector pCGN8100 was transformed into E

coli (strain M15, Qiagen, Hilden, Germany) using a calcium

chloride procedure described in Maniatis, et al. ((1989)

Molecular Cloning: A Laboratory Manual, Second Edition, Cold

Spring Harbor Laboratory, Cold Spring Harbor, New York).

Transformed colonies were screened by Western immunoblot analysis for expression of the HPO Lyase protein using antibodies raised to the bell pepper HPO lyase as described in Shibata, et al. (1995) Plant Cell Physiol. 97:1059-1072.

Hydroperoxide lyase activity was determined by gas chromotography (GC) methods described by Matsui, et al. (1991), Phytochemistry, 30:2109-2113 using linolenic acid 13-hydroperoxide as a substrate.

TABLE 1

		TIDDE I		
Sample	Area	nmole	nmole/10min/mg	
8100 Control	24677	130	153 24	
	4089	28		

The results of the GC analysis shown in Table 1, confirms that the *Arabidopsis* HPO lyase-like sequence encodes a 13-HPO lyase enzyme.

30

5

³⁵ Example 6 Identification of Additional HPO lyase Sequences

10

15

Additional HPO lyase-like sequences are obtained from tomato (Lycopersicon esculentum Mill. Cv UC82B) green immature fruit (ca 20 mm) and cucumber (Cucumis sativus L. cv Suyo) hypocotyl (of 3 day old seedlings) tissues. Total RNA was isolated using TRIzol reagent (Gibco-BRL Life Technologies, Gaithersburg, MD) following the manufacturers protocol.

Complementary DNAs (cDNA) are obtained using the Marathon cDNA Amplification Kit (Clontech, Palo Alto, CA) following the manufacturers directions.

The sequences of HPO lyase from bell pepper (Matsui, et al. (1996) supra), banana (European Patent Application, Publication Number EP 0 801 133 A2) and Arabidopsis were aligned using ClustalW (http://www.clustalw.genome.ac.jp/) and seven conserved peptide sequences were identified (see Figure 7 for positions, Table 2 for a listing).

Table 2

	Peptide Sequence	Primer Name	Oligonucleotide Sequence
1	PGSYG	HPOL1S	5'-ATNCCNGGNWSNTAYGG-3'
2	QPLEEI	HPOL2S	5'-CARCCNYTNGARGARAT-3'
-		HPOL2AS	5'-ATYTCYTCNARNGGYTG-3'
3	GFNAYGG	HPOL3S	5'- GGNTTYAAYGCNTWYGGNGG-3'
		HPOL3AS	5'-CCNCCRSANGCRTTRAANCC-3'
4	YQPLVM	HPOL4S	5'-TAYCARCCNYTNGTNATG-3'
٠.		HPOL4AS	5'-CATNACNARNGGYTGRTA-3'
5	VFDEPE	HPOL5S	5'-GTNTTYGAYGANCCNGA-3'
		HPOL5AS	5'-TCNGGNTCRTCRAANAC-3'
6	NGPQTG	HPOL6AS	5'-CCNGTYTWNGGNCCRTT-3'
7	NKQCAAKD	HPOL7AS	5'-CYTTNGCNGCRCAYTGYTTRTT-3'

A set of synthetic oligonucleotides (Table 2) are synthesized for use in polymerase chain reactions with the cDNAs obtained above to identify sequences which are homologous to HPO lyase sequences. The PCR reactions are carried out using Advantage cDNA Polymerase Mix (Clonetech, Palo Alto,CA) using the reaction conditions according to the manufacturers protocol. The letter "S" in the oligonucleotide name designates a PCR primer designed to amplify the sense strand, or forward reaction primer. The letters "AS"

20

designates a PCR primer designed to amplify the antisense strand, or reverse reaction primer. In the oligonucleotide sequence, the letters "N" represents an A, C, G, or a T, the letter "S" represents a C or a G in that position, the letter "Y" represents a C or a T, and the letter "R" represents an A or a G in that position.

A single PCR product, of approximately 475 bp, was amplifed in reactions containing the primers 4HPOL3S and 11HPOL7AS, with the cDNAs obtained from both cucumber and tomato (described above). The 475 bp PCR product from tomato and cucumber were cloned into the plasmid pCR2.1TOPO (invitrogen) to yield the plasmids T15 (pCGN8305) and C15 (pCGN8309) respectively. In PCR reactions with 6HPOL4S and 11HPOL7AS, a single product, of approximately 200 bp, was obtained from amplification reactions with cDNA obtained from cucumber hypocotyl tissue. The 200 bp product was cloned into pCR2.1 TOPO (Invitrogen), to create the plasmid C17 (pCGN8308).

The nucleotide sequence of each PCR product was

determined by automated sequencing. The sequences obtained are compared to nucleic acid and amino acid sequences of HPO lyase sequences from bell pepper, Arabidopsis, and banana leaf, as well as to DNA and amino acid sequences coding for allene oxide synthases from guayule ((1995) J. Biol. Chem.

25 270(15):8487-8494), flaxseed ((1993) Proc Natl Acad Sci USA 90(18):8519-8523) and Arabidopsis.

The results demonstrate that the T15 nucleic acid sequence is approximately 85% similar to the bell pepper HPO lyase DNA sequence and about 88% similar in the amino acid sequence. Furthermore, the T15 sequence is also at least about 55% similar to other HPO lyase nucleic acid sequences and at least about 57% similar in the amino acid sequence. In addition, the T15 amino acid sequence is only about 41% similar to the allene oxide synthase sequences. The C17 sequence also follows a similar pattern of similarity to the HPO lyase sequences. Thus, the T15 and C17 sequences encode proteins highly similar to HPO lyase.

30

However, the results of the sequence comparisons (Figure 8) demonstrate that the C15 nucleic acid sequence is between 50% and 54% similar to the other HPO lyase nucleic acid sequences and about 58% similar to the allene oxide synthase DNA sequences. Furthermore, the deduced amino acid sequence of C15 is between about 38% and 42% similar to the HPO lyase amino acid sequences and about 51% similar to the AOS amino acid sequences. Thus, the C15 sequence encodes a protein which is divergent from both the known HPO lyase sequences, and is more similar to allene oxide synthase sequences.

The T15, C15 and C17 sequences are used to search Genbank. Search results further confirm the sequences from T15 and C17 as being similar to HPO lyase sequences, while the sequence of C15 is more similar to allene oxide synthase sequences.

In order to obtain full length coding sequence for T15, C15 and C17, RACE PCR reactions are employed using the Marathon cDNA Amplification kit (Clontech) according to the manufacturers protocol, and the oligonucleotides shown in Table 3.

Table 3

1KMC10-1: 5'-CGGTGGAGATCCTCGCCACCGGTGCCGACCC-3'
2KMC10-2: 5'-CTTCCTTCACGGTTGTCCTCACTTCCTCCGCCAG-3'

3KMC17-1: 5'-TCCAGCAGCGCTGCCCCTTTCTCTCCCCGGG-3'
4KMC17-2: 5'-CACTGTTTGTTCTTCTCGCTCGGTGTCCCCG-3'
5KMC10-3: 5'-GGGTCGGCACCGGTGGCGAGGATCTCCACCG-3'
6KMC10-4: 5'-CTGGCGGAGGAAGTGAGGACAACCGTGAAGGAAG-3'
7KMC17-3: 5'-CCGGGGAGAAAAGGGGCAGCGCTGCTGG-3'
30 8KMC17-4: 5'-CGGGGACACCGAGCGAGAAACAAACAGTG-3'
9KMT15-1: 5'-GACTTGGTACTGGTGGACTAAGCCTAAGTGTTTC-3'
10KMT15-2: 5'-GGCTGATAACCACAAAGAAGCTCCCCTTTC-3'
11KMT15-3: 5'-GAAACACTTAGGCTTAGTCCACCAGTACCAAGTC-3'
12KMT15-4: 5'-GAAAGGGGAGCTTCTTTGTGGTTATCAGCC-3'

35

5

10

15

20

PCR products from the amplification reactions with DNA obtained from tomato and cucumber are cloned into pCR2.1 TOPO. The sequences of the 5' and 3'-RACE products from tomato

(pCGN8303 (5' RACE) and pCGN8304 (3' RACE)), cucumber, C15 (pCGN8302 (5'RACE) and pCGN8306 (3' RACE)) and C17 (pCGN8301 (5' RACE) and pCGN8307 (3' RACE)) are sequenced and aligned with the respective sequences obtained from pCGN8305, pCGN8309, and pCGN8308 to obtain preliminary full length sequences corresponding to a tomato HPO lyase-like sequence (Figure 4), a cucumber HPO lyase-like sequence (Figure 5).

10 Example 7 Preparation of Expression Constructs

A set of constructs are prepared for transformation into either plant or bacterial hosts to further characterize the novel sequence from cucumber. To create a full length coding sequence for the cucumber (C15) allene oxide synthase-like sequence, the sequences from the 5' RACE (pCGN8302) and 3' RACE (pCGN8306) were PCR amplified and combined at a unique restriction endonuclease site.

The 5' C15 sequence is amplified using primers (4KMC15ES1 5'- CGGGATCCATGGCTTCTTCCTCCCTGAACTTC-3' and 5KMC15EAS2 5'20 TGCCGACCCATTTCAGTATAGTGGG-3') in PCR amplification reactions described above. The primer 4KMC15EAS1 amplifies from the 5' region and contains the start codon (ATG), and a BamHI site. The 3' C15 sequence is amplified using the AP1 primer provided in the Marathon Kit (BRL-Lifetechnologies, Gaithersburg, MD) and the primer 6KMC15ES3 (5'- TTCACACCATTCCCCTGCCTTTCTCCC-3'). The sequence of the C15 full length clone is shown in Figure 6.

A. Bacterial Expression Construct

The 5' RACE PCR amplification product is digested with

BamHI and XbaI (unique site endogenous to the C15 sequence)
and cloned into the expression vector pQE30 (Invitrogen) with
the amplification product of the 3' RACE PCR reaction digested
with XbaI and SmaI. This construct provides a full length
encoding sequence of the C15 cDNA in the E. coli expression

vector to create the vector pCGN8333. The full length
sequence is also cloned into the plasmid pUC119 to create the
vector pCGN8334.

B. Plant Expression Construct

5

A binary vector for plant transformation, pCGN5138, was constructed from pCGN1558 (McBride and Summerfelt, (1990) Plant Molecular Biology, 14:269-276). The polylinker of pCGN1558 was replaced as an *Hind*III/EcoRI fragment with a polylinker containing unique restriction endonuclease sites, *Hind*III, SseI/PstI, NotI, BamHI, SwaI, XbaI, PacI, AscI, and Asp718.

The full length coding sequence of C15 is cloned to be expressed from the plant constitutive promoter 35S for expression in plants. The expression cassette is cloned into the binary vector pCGN5138 to create the vector pCGN8337.

Example 8 Expression of cucumber C15 in E. coli

The expression vector pCGN8333 was transformed into E coli (strain M15, Qiagen, Hilden, Germany) using a calcium chloride procedure described in Maniatis, et al. ((1989) Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). Transformed colonies were screened by Western immunoblot analysis for expression of the HPO Lyase protein using antibodies raised to the bell pepper HPO lyase as described in Shibata, et al. (1995) Plant Cell Physiol. 97:1059-1072.

Hydroperoxide lyase activity was determined by spectrophotometric and gas chromotography (GC) methods described by Matsui, et al. (1991), Phytochemistry, 30:2109-2113, using both linolenic acid 13-hydroperoxide and linolenic acid 9-hydroperoxide as substrates.

The results of the gas chromotography assay (Figure 9) demonstrate that the protein encoded by the cucumber C15 sequence has greater activity toward linolenic acid 9-hydroperoxide (Figure 9B) substrates than linolenic acid 13-hydroperoxide substrates (Figure 9A). The results of the spectrophotometric assays further demonstrate the preference of the protein encoded by cucumber HPO lyase nucleic acid sequence for 9-Hydroperoxide substrates. The results of the spectrophotometric assay are presented in Figure 10.

10

15

20

25

30

Thus, the cucumber C15 sequence represents the first known cloning of a nucleic acid sequence encoding a 9-hydroperoxide lyase.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claim.

Maria de Cara

Claims

What is Claimed is:

5

- 1. An isolated nucleic acid sequence coding a hydroperoxide lyase with activity toward fatty acid 9-hydroperoxides.
- 2. The sequence according to Claim 1, wherein said sequence has activity toward linolenic acid 9-hydroperoxide.
- 3. The sequence according to Claim 1, wherein said sequence is obtained from a plant source.
- 4. The sequence according to Claim 1, wherein said sequence is obtained from cucumber.
 - 5. The sequence according to Claim 1, wherein said sequence is obtained from cucumber hypocotyl.
- 6. The sequence according to Claim 1, comprising the 15 sequence shown in Figure 6.
 - 7. A construct comprising a promoter functional in a host cell, a sequence encoding a 9-hydroperoxide lyase, and a transcriptional termination sequence.
- 8. A construct according to Claim 7 wherein said 9-20 hydroperoxide lyase sequence is isolated from a plant.
 - 9. A construct according to Claim 7 wherein said 9-hydroperoxide lyase sequence is isolated from cucumber.
 - 10. A method for increasing the resistance of a plant to a plant pathogen comprising expressing an HPO lyase from a construct according to Claim 7, wherein said HPO lysase encoding sequence is linked to DNA sequence capable of directing expression in a plant cell.
- 11. A method for increasing the volatile composition of a plant comprising expressing an HPO lyase from a construct

 30 according to Claim 7, wherein said HPO lysase encoding sequence is linked to DNA sequence capable of directing expression in a plant cell.

ſı,

Δ

Ö

K

S

œ

<u>ت</u>

回

o

Ö

S

[GENETYX: Translation of Nucleotide into Amino Acids]

1997.11.19

: Arabidopsis/lyase/cDNA/truncate Filename

Sequence Position: 1 Size Sequence

ATGGGGGGGACTTCCCCGGGGCCACCACCGTCAACATCCCTAACATCTCAGCAGCCACCATCACCCCCTCACAGCTTCCCCTCGGGATGCCGGGGTT 10

CGTACGCTGGCCGTTGGACCATTATCGGACCGTTTAGATTCCAAGGACCCGATAAGTTTTTCCGGACAAGAGCTGAGAAGTATAAGAGACTCAGAAGTATAAGAGCACTGTG Ö E œ u 4 L o S Д ۵, Ľ. 160 Ŋ Q, O × S

TICCGTACAAATATTCCTCCGACGTTTCCTTTCTTCGGCAACCCATCTTTTGACATGGATCTAGTTGATAAAAGAGATGTTCTCATCGGAGACTTCCCGG Δ œ × Ö > Ω Σ Ω Ĺ, 240 230 220

CCTAGCCTTGGGTTCTACGGCGGCGTTCGTGTTGGAGTTTATCTCGACACTACTGAGCCAAAGCACGCCAAGGTTCGTACATACGTTGCTTTGCTACTAT H K E œ > A K Ħ 回 ۲ 330 310

tacataaaaggtttcgctatggaaacactaaaacgaagctcaaaagtatggctacaagacttcgttcaaacctaaacattttctggggaacaatcgaat z O J = × ſi, S ⋖ œ 440 420

æ

× Ö z 田 н Ω 560 S S 540 520

TGGCTTGCGAGTTATTCCCACTGCTAAA<u>CTTGGCGTAGTTCCTCAGCCTCTTG</u>AAGAGATTTTACTTCATACTTGGCCTTATCCTTCTCTCTTTAA W L A L Q V I P T A K L G V V P Q P L E E I L L H T % P Y P S L L I 640 620

Fig. 16

TTACAGGAGAGGATTAGAACGCAAGGTTTGGGATCGGATCGGATCTTAATTTCAAGACGGTTAACGAAATGGAGCTGGTTAAATCGGTGG L Q E R I R T E V R R V C G S G S D L N F K T V N E M E L V K S V V TCGCCGGAAATTACAAAAAGCTITACAATTCATCGACGAGAATGCCTCCGGTTAGGTCAAGAAGAATTGGGGTTGACCCGAGATGAGGC A G N Y K K L Y N F I D E N A G D C L R L G Q E E F G L T R D E A tatticaaaatcitictcittigittitaggittitaatgcciacggggcittiticcgtcittcittaccitcittigatcggggaataaccggcgacaattccggt G S O R Firri O SLI N F K Œ > CZ CZ 950 940 Bunit Ö O G F 930

TITACGADAÇGCTGCGTTTTAGTCCTCCGGTTCCGCTGCAATTCGCACGTGCGAGGAAAGATTTTCAGATAAGTTCACACGATGCTGTTTTTGAGGTCAA Y E T L R F S P P V P L Q F A R A R K D F Q I S S H D A V F E V K aaaagttaaccttctttgtggttatcagccgctttgtgatgagagacgctaatgttttgacgaacggaggaatttaaaccggaccggtatgtaggtgag 1190 ж Х Ά ь H O 凹 Д F D E 1160 N V 1150 VMRD 1140 1130 O > O 1120 ပ **.** G ×

1070

1040

1030

```
[ GENETYX : Amino Acid Seque .e Homology Data ]
1st Amino Acid Sequence
 File Name
              : Arabidopsis lyase AA
 Sequence Size
               : 492
2nd Amino Acid Sequence
 File Name
             : Bell pepper lyase AA
: 480
 Sequence Size
 Unit Size to compare = 2
[57.0% / 470 aa]
   1 · MLLRTMAATSPRPPPSTSLTSQQPPSPPSQLPLRTMPGSYGWPLVGPLSDRLDYFWFQGP
                       * ** ** ** **** ** *** **** ***
               MIPIMSSAPLSTATPISLPVRKIPGSYGFPLLGPLWDRLDYNWFQKL
  61 · DKFFRTRAEKYKSTVFRTNIPPTFPFFGNVNPNIVAVLDVKSFSHLFDMDLVDKRDVLIG
      48 * PDFFSKRVEKYNSTVFRINVPPCFPFFLGVNPNVVAVLDVKSFAHLFDMEIVEKANVLVG
 121 DFRPSLGFYGGVRVGVYLDITEPKHAKIKGFAMETLKRSSKVWLQELRSNLNIFWGTIES
     108 DFMPSVVYTGDMRVCAYLDTSEPKHTQIKNFSLDILKRSSKIWVPTLVKELDTLFGTFES
181 · EISKNGAASYIFPLORCIFSFLCASLAGVDASVSPDIAENGWKTINTWLALQVIPTAKLG
     168 DLSKSKSASLLPALQKFLFNFFSLTFLGADPSASPEIANSGFAYLDAWLAIQLAPTVSIG
241 · VVPQPLEEILLHTWPYPSLLIAGNYKKLYNFIDENAGDCLRLGQEEFGLTRDEAIQNLLF
     228 * VL-QPLEEIFVHSFSYPYFLVRGGYEKLIKFVKSEAKEVLTRAQTDFQLTEQEATHNLLF
301 · VLGFNAYGGFSVFLPSLIGRITGD-NSGLQERIRTEVR-RVCGSGSDLNFKIVNEMELVK
     287 ILGFNAFGGFTIFLPTLLCNLGDEKNAEMQEKLRKEVREKVGTNQENLSFESVKEMELVQ
359 · SVVYETLRFSPPVPLQFARARKDFQISSHDAVFEVKKGELLCGYQPLVMRDANVFDEPEE
    * *** ** ***** * ****** * ***********
347 * SFVYESLRLSPPVPSQYARARKDFMLSSHDSVYEIKKGELLCGYQPLVMKDPKVFDEPEK
419 FKPDRYVGETGSELLNYLYWSNGPQTGTPSASNKQCAAKDIVTLTASLLVADLFLRYDTI
    407 FMLERFTKEKGKELLNYLFWSNGPQTGSPTESNKQCAAKDAVTLTASLIVAYIFQKYDSV
479' TGDSGSIKAVVKAK
    . .***...* **
467° SFSSGSLTSVKKAC
```

Fig. 6

[GENETYX : Amino Acid Sequence Homology Data]

1st Amino Acid Sequence
 File Name : Arabidopsis lyase AA
 Sequence Size : 492

2nd Amino Acid Sequence
 File Name : Arabidopsis AOS AA
 Sequence Size : 517

Unit Size to compare = 2

[39.2% / 485 aa]

MLLRTMAATSPRPPPSTSLTSQQPPSPPSQLPLRTMPGS 1 MASISTPFPISLHPKTVRSKPLKFRVLTRPIKASGSETPDLTVATRTGSKDLPIRNIPGN 40 · YGWPLVGPLSDRLDYF%FQGPDKFFRTRAEKYKSTVFRTNIPPTFPFFGNVNPNIVAVLD 61 * YGLPIVGPIKDRWDYFYDQGAEFFFKSRIRKYNSTVYRVNMPPG-AFIAE-NPQVVALLD 100 · VKSFSHLFDMDLVDKRDVLIGDFRPSLGFYGGVRVGVYLDTTEPKHAKIKGFAMETLKRS 119 GKSFPVLFDVDKVEKÆLFTGTYMPSTELTGGYRILSYLDPSEPKHEKLKNLLFFLLKSS 160 · SKVWLQELRSNLNIFWGTIESEISKNGAASYIFPLQRCIFSFLCASLAGVDASVSPDIAE ************ 179 RNRIFPEFQATYSELFDSLEKEAFPLRESGFRRFQRRNRLLFLGSSFL-RDESRRYKLKA 220' NGWKTINIWLALQVIPTAKLGVVPQPLEEILLHTWPYPSLLIAGNYKKLYNFIDENAGDC 238 DAPGLITKWVLFNLHPLLSIG-LPRVIEEPLIHTFSLPPALVKSDYQRLYEFLRIR-GEI 280 · LRLGQEEFGLTRDEAICMLLFVLGFNAYGGFSVFLPSLIGRITGDNSGLQERIRTEVRRV 296 L-VEADKLGISREEATENLLFATSFNTWGGMKILFPNMVKRIGPGGHQVHNRLAEEIRSV 340 · C-GSGSDLNFKTVNEMELVKSVVYETLRFSPPVPLQFARARKDFQISSHDAVFEVKKGEL 355 " IKSNGGELTMGAIEKMETTKSVVYECLRFEPPVTAQYGRAKKOLVIESHDAAFKVKAGEM 399 · LCGYQPLVMRDANVFDEPEEFKPDRYVGETGSELLNYLYWSNGPQTGTPSASNKQCAAKD 415 LYGYQPLATRDPKIFDRADEFVPERFVGEEGEKLLRHVLWSNGPETETPTVGNKQCAGKD 459 · IVTLTASLLVADLFLRYDTITGDSGSIKAVVKAK .*.*.*.*. . . * ***. . . *. 475 FVVLVARLFVIEIFRRYDSFDIEVGTSPLGSSVNFSSLRKASF

Fig. 3

```
[GENETYX: Translation of Nucleotide into Amino Acids]
  1998.06.24
  Filename
             : kmt15-1/full copy
  Sequence Size
             : 1901
  Sequence Position: 2 - 1901
       10
              20
                    30
  40
             120
 130
 230
                           MNSAPLSTPAPVTLPVRSIP
 nggcagctacggattgccgttagtagggccaatcgcggatcgattagactacttctggttccaaaaaccggaaaacttcttcaccaagagaatggaaaag
                   330
  G S Y G L P L V G P I A D R L D Y F W F Q K P E N F F T K R M E K
                  430
 cacaagagcacggttttcagaacgaacgtgccgccgtgttttccgtttttcggtagtgtgaatccaaatgtggtggcggttttagatgtgaaatcgtttt
 H K S T V F R T N V P P C F P F F G S V N P N V V A V L D V K S F S
 cgcatctgtttgatatggagattgtagagaaagctaatgtgcttgttggggatttcatgcccagtgttgtttatactggagatatgcgtgtttgtgctta
                  530
  H L F D M E I V E K A N V L V G D F M P S V V Y T G D M R V C A Y
 tctcgatacttctgaacctaaacatgcccagattaagaatttttcacaggatattctaaaaagaggctcaaaaacatgggtgcctacactacttaaagag
                 630
                        640 650
 LDTSEPKHAQIKNFSQDILKRGSKTWVPTLLKE
            720
                730
cttgatacaatgtttacaacttttgaagcagatctttcaaaatccaatacagcttctcttcttcctgcactccaaaaattcctctttaacttcttctccc
                        740 750
L D T M F T T F E A D L S K S N T A S L L P A L Q K F L F N F F S L
           820
                       840 850
830
 TILGADPSVSPEIANSGYIFLDSWLAIQLAPTV
    910 920 930
                      940
cagcattggtgtccttcaacctcttgaagaaattttagtccattcttttgcataccctttttttcttgtcaaaggtaattatgaaaaactcgttcaattc
 SIGVLQPLEEILVHSFAYPFFLVKGNYEKLVQF
          1020 1030 1040
gtcaaaaatgaagccaaggaagttttaagcagggcacaaacagagtttcaactcacagaacaagaagccattcataaccttttgttcattcttgggttca
V K N E A K E V L S R A Q T E F Q L T E Q E A I H N L L F I L G F N
                      1140 1150
atgettttggtggettetecattttettgccaactettttgggaaatettggagatgagaaaacgcggatatgcaagagaaactgagaaaagtgag
                 1130
 AFGGFSIFLPTLLGNLGDEKNADMQEKLRKEVR
          1220
                1230
agacaaagtcggcgtaaatccagaaaacttgagttttgaaagtgttaaagaaatggaacttgttcagtcttttgtttatgaaacacttaggcttagtcca
D K V G V N P E N L S F E S V K E M E L V Q S F V Y E T L R L S P
          1320
ccagtaccaagtcaatatgcaagagcaagaaaagattttaaactgagttcacatgattcagtttacgaaatcaagaaaggggagcttctttgtggttatc
                 1330
P V P S Q Y A R A R K D F K L S S H D S V Y E I K K G E L L C G Y Q
                        Fig. 4A
```

agcctttagttatgaaagatccaaaggtgtttgatgaacctgaaaagtttgtgttggagagatttacaaaggaaaaagggaaagaattgctgaattattt PLVMKDPKVFDEPEKFVLERFTKEKGKELLNYL gttttggtctaatggtccacaaactgggagacctactgaatcaaacaagcaatgtgctgctaaagatatggttactctaactgcttctttgattgtggct FWSNGPQTGRPTESNKQCAAKDMVTLTAS,LIVA Y I F Q K Y D S V S F S S G S L T S V K K A S *

[GENETYX: Translation of Nucleotide into Amino Acids]

1998.06.24

Filename

Filename : C17/full copy Sequence Size : 1712 Sequence Position: 1 - 1712

								•	
10	20	30							
ctcctctctctctctctctctctctctctctctctctc	ctaccggas	aagttcaacc	40	50	60	· 70	80	00	
				tecetttea	aattcatacto	ccacatete	Caaatatona	tacacata	100
110	120	130	140					rycacccca	catgacct
ccggcggaatgc	cttccatac	cttcatctat	ttccccccc.	150	160	170	180	190	
				caaccccff	tacctctcaga	aatatccccg	gcagctacon	tttaccan-	200
210	220	230	240	25.0				geegeeg	ctgggatc
categgtgaceg	actcgatta	ctactggttt	aaggacccga	250	260	270	280	290	
categgtgaceg			55 -	-HAG CCCCC	caggactcgta	tggagaagaa	tcgaagcacg	gttttcaraa	300
310	320	330	3,40	250			-	J	ccaacgcc
cctccgtctttc	cctttcatc	teegetgatee	 	350	360	370	380	390	400
cctccgtctttc		- 0	gag cag c cg	caarrerad	yattgcaagtc	ctttgcgcat	Ctatttgaca		400
410	420	430	440	450				-99aaaccgc	ggagaaaa
ataatgttctgg	teggtgatti	tcatgcccago	atcagtttcac	450	460	470	480	400	
ataatgttctgg		_	succae	cygaaacat	gagagtgtgt	gcttatttgg	atacctcggaz	NCCCARTORS	500
510	520	530	540	550	0 _0		- 33		cccaaggt
adagaacttcat	tacagacat	ctacggcgga	gctcaaggara	tagatata	560	570	580	590	
aaagaacttcat				cagacaccg	gagttggaat	ccaacctatc	gacgatgtggg	acaacatte.	600
910000000	620	630	640	650				-39 -44 - 69	actagaa
araacaaaaaaca	aacaatcgg	gtttcagaaa	ctttctccaac		650	670	680	690	700
atggcaaagaaca				3	ccaatttcttc	tgrazaacto	tegeeggege	cgacactocs	700
COGRAGE	720	730	740	750	244				mad ccac
-ssung cygcaaa	acccggata	catagacgte	ataatctggcto		760	770	780	790	900
cggaagtggcaaa 810	000				accartccc3	ccatccacat	cggcgttctc	cagcccctto	ARGEDES
attectteactet	820	830	840	850	860			•	
	.ccccgetea	ccctttttcc	cattgetteed	gctaccaa	agactctaces	870	880	890	900
attectteactet	620				-saccetacya	CCCCCCCAA	aaagaagggg	cagaagttgt	tgagega
ggtgtcaccgage	320	930	940	950	960	870			
os systemategage	cegggetaa	caaaggaagaa	gccattcacaa	tctcatct	Caccata	370	980	990	1000
ggtgtcaccgagt	1000		•	4.41	-uuccatgggc	ttcaacgcct	atggtggcttd	agtetette	ttcccta
ttctacttgaccg	1020	1030	1040	1050	1060	1000			
-3	cucactcaat	rgataaaaccg	gtctacaacaa	agaateete	Baggaagtca	10/0	1080	1090	1100
ttctacttgaccg	1120	1400	-		-33449664	AA CCAAAA CC	cagetetggte	tgaccttcg	agtcagt
caaggagatggac	Etcatorace	1130	1140	1150	1160	1170		•	
caaggagatggacd		-ccaccgceta	cgagacactcc	ggcttgacc	cccagttcc	+1/U	1180	1190	1200
1210	1220	1220				·	ccagagecag	aagatttca	agctaac
ttcctacaattcaa	COLATARCA	1230	1240	1250	1260	1270	1000		
		cedayaaagg	ggaactgcttt	gtgggtatc	agccgctggto	atgagagae	1280	1290	1300
ttcctacaattcaa	1320	1330	1242				cayaggtgtt	cgatgaaccg	rgaggeg
tttaatccagaccg	atteegggg	agagaaagg	1340	1350	1360	1370	1300		
tttaatccagaccg		3 3 4 4 4 4 4 4	acadedeedeef	gattattt	attttggtcga	atgggccgca	22000	1390	1400
1410	1420	1430	1440				aacggggaca.	ccgagcgaga	agaaca
aacagtgtgctggg	aaggattta	gtagtactas	1440	1450	1460	1470	1480	1400	
aacagtgtgctggg			.aaaaaacaacac	Etgtggctt	acatatttaa	aaggtatgat	tegattgera	1490	1500
1510	1520	1530	1540			•	gauge eg	ayyaaggagg	ttccat
tacagcttttcaaa	gggctaatt	gaaaagtgaaa	1340	1550	1560	1570	1580	7500	
tacagcttttcaaa 1610		2-3-40	-a-aaarcaca6	acacattat	gtatcgttta	ttgagagatg	cactttter	137U	1600
1610	1620	1630	1640	1650				caacggc	EECEEE
ccacgtatgagtt	tggagccaaa	atgggaagat	ttggaargase	T030.	1660	1670	1680	1690	1700
cttatgtatgagtt			J J - S L Y CAL	uaa ccaata	aaatcaagatt	tactttacct	tgtaaaaaaa	.0.00	1/00
1/10			_						aaaaa
aaaaaaaaaa			Tio E		•				
			Fig. 5	•	1				
			J		•				

[GENETYX: Translation of NucleoLide into Amino Acids]

1998.06.24

Filename

: C15 full/cDNA copy

Sequence Size : 1813 Sequence Position: 2 - 1813

10 20	30 40	50 60 .	70 80	90 100
ccattctcttccaacgtgaagataag	JAMACCECELGECTACECECEG	cacgatcacaggtcacagc	aatggettetteeteeet	gaacttcctctcaaa
			MASSSP	ELPLK
110 120		150 160	170 180	190 200
cccattcccggtggctatggcttccc	cttcctcggtcccatcaaaga	ccgttacgattacttctat	ttccaaggtagagacgaat	tettecotteeraa
PIPGGYGFP	FLGPIKD	RYDYFY	FQGRDEF	FRSRI
210 220	230 240	350 360	222	22.
ttaccaaatacaactccaccgtcttc		250 260 cccttcatctcctccgatt	270 280	290 300
TKYNSTVF	HANMPPG	P F I S S D S	R V V V t. T.	D A T. C E
				-
310 320		350 360	370 380	390 400
tcccatcctcttcgacaccaccaaag	gtcgagaaacgcaacattctcg	acggaacttacatgccctc	cttgtccttcaccggcgg	attogoadotgtgot
PILFDTTK	VEKRNIL D	GTYMPS	LSFTGG	IRTCA
410 420	430 440	450 460	470 480	490 500
tatttggacccatcggaaacagagca	acactgttctcaaacgcctctt	teteteetttetegettet	caccatgacaggttcatc	ctctgtttcgaager
YLDPSETEH	TVLKRLF	L S F L A S	HHDRFII	PLFRSS
510 520	530			· ·
		550 560	570 580	590 600
ccttgtctgagatgtttgttaagett L S E M F V K L	EDKLADK	N K I A D F N	actegattagtgatgeegt 1 S T S D 1 V	gregitigatiatgi
	:	7		3
		650 660	670 680	690 700
tttccgtttattctccgatggaacc	cctgattcgacattagctgctg	atggacctggaatgttcg	atttatggcttgggcttca	acttgccccattggct
FRLFSDGT		GPGMFD	r w r c r o	LAPLA
710 720	730 740	750 760	770 780	790 800
tccattggccttcccaaaattttctc	ctgtttttgaagatctcattat	tcacaccattcccctgcct	ttcttcccagtcaagagt	cottacaggaagerrr
SIGLPKIFS	V F E D L I I	H T I P L P	FFPVKSI	RYRKLY
810 820	830 840	252		
ataaagcgttttactcctctctgg		850 860 3885899999358935899	870 880	890 900
KAFYSSSG	SFLDEAE	K Q G I D R I	E K A C H N L	V F L A G
910 920		950 960	970 980	99 0 1000
attcaacgcatacgggggaatgaaa F N A Y G G M K 1	gtcctttttcccactatactga V 1 F b m 1 1 V	aatgggteggeaeeggtgg	gcgaggatetecacegtaa	actggcggaggaagtg
		. * * * * * * * * *	E D L H R K	LAEEV
1010 1020	1030 1040 1	050 1060	1070 1080	1090 1100
aggacaaccgtgaaggaagaagggg	gactgactttctccgccttgga	gaaaatgagtctgctgaag	gtcggtcgtgtacgaagct	Ctgaggatcgaaccgc
RTTVKEEGG	LTFSALE	K M S L L K	SVVYEA	LRIEPP
1110 1120	1130 1140 1	150 1160	1170 1100	
cggtgccgttccagtacgggaaagc			1170 1180	1190 1200
V P F Q Y G K A	KEDIVIQ	SHDSCF	K I K K G E T	I F G Y Q
1210 1220	1230 1240 1	250 1260	1270 1280	1290 1300
gccgtttgctactaaagatccgaag	atttttaaggactc gg agaagt	tcgtgggcgataggttcg	Laggagaagaagagaga	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
P F A T K D P K	IFKDSEKF	VGDRFV	G E E G E K	LLKYV
1310 1320	1330 1340 1	.350 1360	1370 1300	1200
tactggtcaaatgagcgggagacgg	tggagccgacggcggagaacaa	gcagtgtccggggaagaa	1370 1380 Ectggtggtgatgatgggg	1390 1400
YWSNERETV	EPTAENK	QCPGKN	L V V M M G	R I I V V E
	Fig. 4	, A		

1410 aattetteeteegttatgatacgtteactgtggacgtegcagatttggegetgggeeggeggtgaagtteaagteettgaccagagecaccgetteggt 1430 FFLRYDTFTVDVADLALGPAVKFKSLTRATAS 1520 1530 1540 1630 acatgacacatcattgatttgttggggttataaagtaatttcctcgtgatccgcgtcgtttttaatttgagatctcattgtgtgttgtaacccgcgggtg 1720 1730 1740 1810 aaaaaaaaaaa

Fig. 6B

CLUSTAL W (1.7) Multiple Sequence Alignments

```
Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: gil1272340
Sequence 2: banana
                                480 aa
                               483 aa
Sequence 3: ARABIDOPSIS
Start of Pairwise alignments
                                492 aa
Aligning..
Sequences (1:2) Aligned. Score:
Sequences (1:3) Aligned. Score:
Sequences (2:3) Aligned. Score:
                                    51
                                    55
Start of Multiple Alignment
There are 2 groups
Aligning...
Group 1: Sequences:
Group 2: Sequences:
                               Score: 4874
                               Score: 4967
Alignment Score 5148
CLUSTAL-Alignment file created [clustalw.aln]
CLUSTAL W (1.7) multiple sequence alignment
gi|1272340
                  -----MIPIMSSAPLSTATPISLPVRKIPGSYGFPLLGPLWDRLDYNWFO
                  ------MAMMWSSASATAVT--TLPTRPIPGSYGPPLVGPLKDRLDYFWFQ
banana
                  MLLRTMAATSPRPPPSTSLTSQQPPSPPS--QLPLRT MPGSYGWPLVGPLSDRLDYFWFQ
ARABIDOPSIS
                                      : *. . :. :
                                                     ** * : **** **: *** ****
                  KLPDFFSKRVEKYNSTVFRTNVPPCFPFFLGVNPNVVAVLDVKSFAHLFDMEIVEKANVL
GPETFFRSRMATHKSTVFRTNMPPTFPFFVGVDPRVVTVLDCTSFSALFDLQVVEKKNIL
gi|1272340
banana
ARABIDOPSIS
                  GPDKFFRTRAEKYKSTVFRTNIPPTFPFFGNVNPNIVAVLDVKSFSHLFDMDLVDKRDVL
                              .::*******:** **** .*:*:*** .**: ***:::*:*
                  VGDFMPSVVYTGDMRVCAYLDTSEPKHTQIKNFSLDILKRSSKTWVPTLVKELDTLFGTF
gi|1272340
                  IGDYMPSLSFTGDTRVVVYLDPSEPDHARVKSFCLQLLRRGAKTWVSSFLSNLDVMLATI
banana
ARABIDOPSIS
                  IGDFRPSLGFYGGVRVGVYLDTTEPKHAKIKGFAMETLKRSSKVWLQELRSNLNIFWGTI
                  :**: **: : *. ** .***.:*.*:: *:*: *:*:
gi|1272340
                  ESDLSKSKSASLLPALQKFLFNFFSLTFLGADPSASPEIANSGFAYLDAWLAIQLAPTVS
banana
                  EQGIAKDGSAGLFGPLQKCIFAFLCKSIIGADPSVSPDVGENGFVMLDKWLALQLLPTVK
ARABIDOPSIS
                  ESEISKNGAASYIFPLQRCIFSFLCASLAGVDASVSPDIAENGWKTINTWLALQVIPTAK
                  *. ::*. :*. : .**: :* *:. :: *.*.*.**::.:.*;
gi | 1272340
                  IG-VLQPLEEIFVHSFSYPYFLVRGGYEKLIKFVKSEAKEVLTRAQTDFQLTEQEAIHNL
                  VGAIPOPLEEILLHSFPLPFFLVSRDYRKLYEFVEKOGQEVVRRAETEHGLSKHDAINNI
banana
                  ARABIDOPSIS
                  LFILGFNAFGGFTIFLPTLLGNLGDEKNAEMQEKLRKEVREKVGTNQEN-LSFESVKEME
LFVLGFNAFGGFSVFFPTLLTTIGRDK-TGLREKLKDEVRRVMKSRGEKRPSFETVREME
gi|1272340
banana
                  ARABIDOPSIS
gi|1272340
                  LVRSTVYEVLRLNPPVPLQUGRARTDFTLNSHDAAFKVEKGELLCGYQPLVMRDPAVFDD
banana
                  LVRSTVYEVLRLNPPVPLQUGRARTDFILMSDDAGERVLRGELLCG YOPLVMRDANVFDE LVKSVVYETLRFSPPVPLQFARARKDFQISSHDAVFEVKKGELLCG YOPLVMRDANVFDE
ARABIDOPSIS
                  PEKFMLERFTKEKGKELLNYLFWSNGPQTGSPTESNKQCAAKDAVTLTASLIVAYIFQKY
gi11272340
banana
                  PETFAPERFMGS-GKELLKYVFWSNGPETGTPTPANKQCAAKDYVVETACLLMAEIFYRY
                  PEEFKPDRYVGETGSELLNYLYWSNGPOTGTPSASNKOCAAKDIVTLTASLLVADLFLRY
ARABIDOPSIS
gi|1272340
                                                        11 HPOL 7AB
                  DSVSFSSGSLTSVKKAC----
banana
                  DEFVCADDAISVTKLDRAREWE
ARABIDOPSIS
                  DTITGDSGSIKAVVKAK----
                        ..::. .
```

Fig. 7

1. nucleotide

		Γ	_ S	T	9.		_	1	IC.	2	7	-	7	<u>.</u>	œ	2]	α	श	_
			A AO														5		
	,	2	200	7 03	32.4	000	48.8		54.5		52.1		56.4		28.				62.25
		ב ה		50.0	2:33	50 0	2	20.03	36.3	600	32.0		24.1			0 0 2	50.3	67.0	0./0
				55.0	l	55.6	l		I	25	2:3			53.0		20.0	5.00	ביב	0
	10001		13	29.7	10.01	C.RC	2	0.10				583		52.1		52.31		21.7	
	C17	5	200	00.0	F0 7	23.1			0.10	5	7 62	28.7	000	22.8	2 73	0.4.0	5	22.5	
	Le HPOI		84 6				20 5		20 2	2	SE C	0.50	507	7.3	40.7		520	7.00	
100	Carrol			3	04.0	3	2.00		59.7		55.1		20 0		52.1		52.5		
		Campo	Sa LIF OF	PHDO	10.	Cath HDO!	30 11 05	יים ועסרו וע	אס שוני וכ	LOGO SY	JOLUSIA JOLUSIA	1001127	JOHN CIST	000	200	004	ACS		

rigille 8A

2. amino acid

			AI AOS		40.0	7 9	40.4	1		000	38.8	1000	43.2	217	F	60	3	58.2	
		ļ.	LIAOS	00,		40 5		410		11 17		44 5	2.5	512		62.9			
		000	CON NO	410		41.6		43.3		41		44.5		51.6		, ,,,	0 00	60.6	
		7	2	39 6		39.4	1.5	41./	0	35.55	ľ	41.9	İ		212	2.50	512	3	
•		Ms HPO		54.	3.5	5/.3	F2 G	0.50	צט	200			110	21.3	44 5		44.5		
	1001114		21.5	0.70	5 A 7	700.7	56.2				- AR	3	38 4		-		41.1	900	
	23.7	5	50 E	0.20	. 614			000	200.		53.6		7.17	007	10.0	-	11.3	1 1 K	
	Le HPOI	3	88 2			7 10	7.	50 7	700.	F7.0	ى. كن ب	7 00	4.60	416		40.5		40.4	
	Ca HPOL Le HPOI			1000	7.00	59.5	2	57.0		24 1	5	30 8	3	41.0		40.3	199	40.0	
		Caupol	Caliron	P HDO		Cs17 HPOL	.00:	A HPOL		MS HPOL		CS15 HPOL		au AOS	000	200	AOS II	2000	

Floure AR

Flaure 8

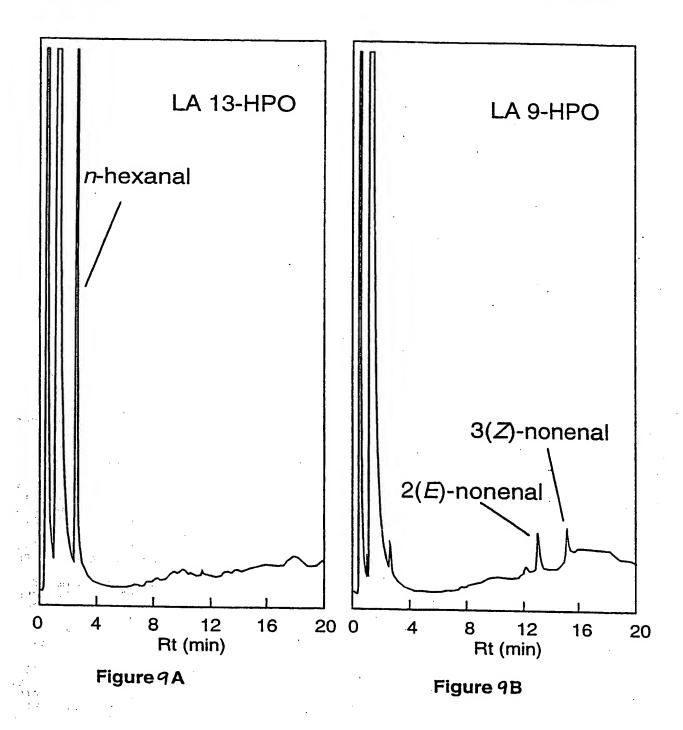


Figure 9

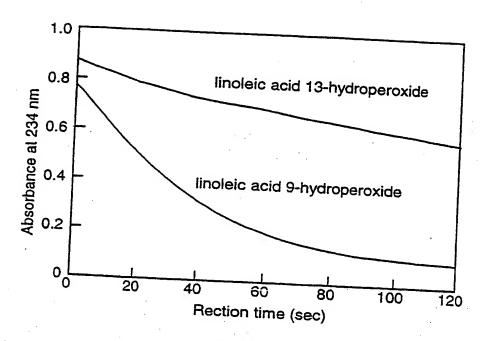


Figure 10

SEQUENCE LISTING

```
<110> Matsui, Kenji
<120> Fatty Acid 9-hydroperoxide Lyase Nucleic
  Acid Sequences
<130> 15457/00/WO
<140> New Application
<141> 2000-02-25
<160> 44
<170> FastSEO for Windows Version 4.0
<210> 1
<211> 1385
<212> DNA
<213> Arabidopsis
<400> 1
atggcggcga cttccccgcg gccaccaccg tcaacatccc taacatctca gcagccacca
                                                                                            60
tcaccccct cacagettec ceteegtaca atgeegggat egtaeggetg geegttggtt
                                                                                           120
ggaccattat cggaccgttt agattccaag gacccgataa gtttttccgg acaagagctg
                                                                                           180
agaagtataa gagcactgtg ttccgtacaa atattcctcc gacgtttcct ttcttcggca
                                                                                           240
acceatettt ttgacatgga tetagttgat aaaagagatg tteteategg agaetteegg
                                                                                           300
cctagccttg ggttctacgg cggcgttcgt gttggagttt atctcgacac tactgagcca aagcacgcca aggttcgtac atacgttgct ttgctactat tacataaaag gtttcgctat
                                                                                           360
                                                                                           420
ggaaacacta aaacgaagct caaaagtatg gctacaagag cttcgttcaa acctaaacat
                                                                                           480
tttctgggga acaatcgaat ccgaaatctc caaaaacggt tcctctgcgc ctctctcgcc.
                                                                                           540
ggcgttgacg cttccgtatc gccggacatc gctgagaacg gttggaaaac aatcaatact tggcttgcgt tgcaagttat tcccactgct aaacttggcg tagttcctca gcctcttgaa gagattttac ttcatacttg gccttatcct tctctcttaa tcgccggaaa ttacaaaaag
                                                                                           600
                                                                                           660
                                                                                           720
ctttacaatt tcatcgacga gaacgccgga gattgtctcc ggttaggtca agaagaattc
                                                                                           780
gggttgaccc gagatgaggc tattcaaaat cttctctttg ttttaggttt taatgcctac gggggctttt ccgtcttctt accttctttg atcgggagaa taaccggcga caattccggt
                                                                                           840
                                                                                           900
ttacaggaga ggattagaac cgaagtcagg agagtttgcg gatccgggtc ggatcttaat
                                                                                           960
ttcaagacgg ttaacgaaat ggagctggtt aaatccgtgg tttacgaaac gctgcgtttt
                                                                                         1020
agtcctccgg ttccgctgca attcgcacgt gcgaggaaag attttcagat aagttcacac gatgctgttt ttgaggtcaa gaaaggtgag cttctttgtg gttatcagcc gcttgtgatg agagacgcta atgtttttga cgaaccggag gaatttaaac cggaccggta tgtaggtgag
                                                                                         1080
                                                                                         1140
                                                                                         1200
accgggtctg aattgctgaa ttatctctac tggtctaacg gtccacaaac cggtaccccg
                                                                                         1260
agcgcgtcga acaaacagtg tgcagctaag gacattgtca ctctcacggc ttccttgctc
                                                                                         1320
gttgccgatt tatttctccg atcaaagact tcaaaaactg aaacacaaca accatcatcg
                                                                                         1380
tatgg
                                                                                         1385
<210> 2
<211> 462
<212> PRT
<213> Arabidopsis
```

- 1..-

Met Ala Ala Thr Ser Pro Arg Pro Pro Pro Ser Thr Ser Leu Thr Ser

<400> 2

1					5															
G1	n Gl	n P	ro	Pro	Se	r Pr	O D.	ro c	·~~		_ 1	0	_					15	i	
G1	v Se	ጉ ጥ	(<i>27</i> ~	20 Gly	· m		υ ₋ .		er.	25	nь	eu	Pro) Le	u Aı	g	Th	r Me	t Pr	co
50	,	3	5	- -	TI:	p Pr	о Ге	eu V 4	al 0	. G1	у P:	ro	Let	ı Se	r As	q	Ar	g Le	u As	q
	50	S A	sp	Pro	Ile	e Se	r Pl 55	ne S	er	G1	y G	ln	Glu	Le	u Ar	g	Se:	r Il	e Ar	g
65	а ње	u Cy	/S	Ser	Va.	l G1:	n II	le P	he	Le	u A	rg	Arg	r Ph	e Le	u	Se	r Se	r Al	a
TI	r Hi	s Le	eu :	Phe	Asp) Me	t As	p L	eu	Va.	1 As	g	Lys	Arc	g As	σ,	Va:	l Le	80 11 Tl)
Gl	/ As	p Pł	ıe i	Arg	Pro	Se	r Le	u G	1y	Phe	e Ty	r	Gly	Gly	y Va	1 2	Arc	95 Va	1 G1	v
Va]	L Ty	r Le	u A	Asp	Thr	Th	c Gl	u P	ro	10: Lys	5 s Hi	.s	Ala	Lvs	s Va	1 3	11(Arc) T Th	- от r Ту	
Val	Al	a Le	u I	Ŀeu	Leu	Let	ı Hi	1: s L	20 Ys	Arg	g Ph	ıe	Ara	ው ጥህን	12	5	A c =	mb.	r Ly	T
Thr	Ly	5 Le	u I	ys	Ser	Met	13 : Al	5 a T]	hr	Arc	- r Al	a	Ser	14()		n		r Ly:	S
Phe	Lei	1 Gl	y A	sn	Asn	150) 7 Il	e A1	ra	λετ	,		155	PHE	: гу:	S 1	Pro	ь Гу:	5 Hi: 160 1 Cy:	s 0
Ala	Ser	Le	u A	la	165 Glv	Val	λ	n 2	- 9	Al-	17	0	GIN	гуз	Ar	g I	?he	Let 175	ı Су: 5	s
Glu	Asr	ı Gl	1 งา	80	Laze	Ψh~	· 713	o As	υ	185	i se	r	Val	Ser	Pro) A 1	\sp 90	Ile	Ala	a
Pro	Thr	19 Al	5 5	-r	Lou	23		2 AS	on O	Thr	Tr	p	Leu	Ala	Let 205	1 G	ln	Va]	l Ile	⊋
Len	210 His	Th	~ m		neu	GIA	21:	L Va 5	1	Pro	G1:	n	Pro	Leu 220	Gli	1 G	lu	Ile	Leu	1
225	7	, III	r T	rp	Pro	Tyr 230	Pro	S∈	r	Leu	Le	u	Ile 235	Ala	Gly	, A	sn	Тух	Lys	3
ole:	reu	1 y :	r A	sn	Phe 245	Ile	Ası) G1	u	Asn	Ala 25	a (Gly	Asp	Суз	L	eu	Arg	240 Leu)
- GTÀ	GIN	Gl	u G 2	lu 60	Phe	Gly	Let	Th	r	Arg 265	Ası)) (Glu	Ala	Ile	G	ln	255 Asn	Leu	L
ren	Phe	Va.	L L	eu	Gly	Phe	Asr	Al	a	Tyr	G13	7 (Gly	Phe	Ser	V	70 al	Phe	Len	
Pro	Ser 290	Let	1 I	le	Gly	Arg	Ile	Th	r	G1y	Asp	> 2	Asn	Ser	Gly	L	eu	Gln	Gĺu	1
305	TTE	Arg	J T	hr (Glu	Val	Arg	Ar	g '	Val	Cys	3 (3ly:	Ser	Glv	S	er	Asn	Len	
Asn	Phe	Lys	T	ır !	Val	310 Asn	Glu	Me	t (Glu	Leu	. J	315 /al	Lvs	Ser	V:	a 7	Val	320 Tyr	
Glu	Thr	Lev	ı Aı	g	Phe	Ser	Pro	Pro	o 1	Val	330 Pro)): I	.eu	Gln	Pho	ν.	1 2	335	Ala	
Arg	Lys	Asp	Pł	ie (Gln	Ile	Ser	Se	r I	345 His	Asp	· A	la:	∵i Va1	Dho	35	50	nry	Ald	
Lys	Gly	Glu	L∈	eu I	Leu	Cys	Gly	360 Tv:) - (Sln	Pro	: . T	u:	701	365	· ·	Lu	vaı	гљз	
Asn 385	Val	Phe	As	g q	3lu	Pro	375 Glu	Gli	, E)he	Tura	. ~	eu 	380	met.	A1	rg	Asp	Ala	
385 Glu	Thr	Gly	Se	ro	lu	390 Leu	T.eu	200		·	TA Z	3	95	Asp	Arg	Т	T	Val	Gly 400	
Glu Gln	Thr	Glv	ፓ ኮ	r F	105	Ser	Δ3-	na!		УĽ	410	т	yr : -	ľrp	Ser	As	sn	Gly 415	Pro	
Gln Ile	Val	ጥኮ ~	42	0	ib-	71- PET	urg	ser	- A	lsn 125	ГЛS	G	ln (Cys .	Ala	A1 43	.a	Lys	Asp	
Ile		435	ne	u 1	TIE.	wig	ser	Leu 440	L	eu	Val	A	la 2	\sp	Leu 445	Ph	ie	Leu	Arg	

Ser Lys Thr Ser Lys Thr Glu Thr Gln Gln Pro Ser Ser Tyr <210> 3 <211> 492 <212> PRT <213> Arabidopsis <400> 3 Met Leu Leu Arg Thr Met Ala Ala Thr Ser Pro Arg Pro Pro Pro Ser Thr Ser Leu Thr Ser Gln Gln Pro Pro Ser Pro Pro Ser Gln Leu Pro Leu Arg Thr Met Pro Gly Ser Tyr Gly Trp Pro Leu Val Gly Pro Leu Ser Asp Arg Leu Asp Tyr Phe Trp Phe Gln Gly Pro Asp Lys Phe Phe Arg Thr Arg Ala Glu Lys Tyr Lys Ser Thr Val Phe Arg Thr Asn Ile Pro Pro Thr Phe Pro Phe Phe Gly Asn Val Asn Pro Asn Ile Val Ala Val Leu Asp Val Lys Ser Phe Ser His Leu Phe Asp Met Asp Leu Val Asp Lys Arg Asp Val Leu Ile Gly Asp Phe Arg Pro Ser Leu Gly Phe Tyr Gly Gly Val Arg Val Gly Val Tyr Leu Asp Thr Thr Glu Pro Lys His Ala Lys Ile Lys Gly Phe Ala Met Glu Thr Leu Lys Arg Ser Ser Lys Val Trp Leu Gln Glu Leu Arg Ser Asn Leu Asn Ile Phe Trp Gly Thr Ile Glu Ser Glu Ile Ser Lys Asn Gly Ala Ala Ser Tyr Ile Phe Pro Leu Gln Arg Cys Ile Phe Ser Phe Leu Cys Ala Ser Leu Ala Gly Val Asp Ala Ser Val Ser Pro Asp Ile Ala Glu Asn Gly Trp Lys Thr Ile Asn Thr Trp Leu Ala Leu Gln Val Ile Pro Thr Ala Lys Leu Gly Val Val Pro Gln Pro Leu Glu Glu Ile Leu Leu His Thr Trp Pro Tyr Pro Ser Leu Leu Ile Ala Gly Asn Tyr Lys Lys Leu Tyr Asn Phe Ile Asp Glu Asn Ala Gly Asp Cys Leu Arg Leu Gly Gln Glu Glu Phe Gly Leu Thr Arg Asp Glu Ala Ile Gln Asn Leu Leu Phe Val Leu Gly Phe Asn Ala Tyr Gly Gly Phe Ser Val Phe Leu Pro Ser Leu Ile Gly Arg Ile Thr Gly Asp Asn Ser Gly Leu Gln Glu Arg Ile Arg Thr Glu Val Arg Arg Val Cys Gly Ser Gly Ser Asp Leu Asn Phe Lys Thr Val Asn Glu Met Glu Leu Val Lys Ser Val Val Tyr Glu Thr Leu Arg Phe Ser

```
Pro Pro Val Pro Leu Gln Phe Ala Arg Ala Arg Lys Asp Phe Gln Ile
Ser Ser His Asp Ala Val Phe Glu Val Lys Lys Gly Glu Leu Leu Cys
Gly Tyr Gln Pro Leu Val Met Arg Asp Ala Asn Val Phe Asp Glu Pro
                                         395
Glu Glu Phe Lys Pro Asp Arg Tyr Val Gly Glu Thr Gly Ser Glu Leu
                                     410
Leu Asn Tyr Leu Tyr Trp Ser Asn Gly Pro Gln Thr Gly Thr Pro Ser
                                425
Ala Ser Asn Lys Gln Cys Ala Ala Lys Asp Ile Val Thr Leu Thr Ala
Ser Leu Leu Val Ala Asp Leu Phe Leu Arg Tyr Asp Thr Ile Thr Gly
Asp Ser Gly Ser Ile Lys Ala Val Val Lys Ala Lys
                                        475
                485
<210> 4
<211> 480
<212> PRT
<213> Bell Pepper
```

<400> 4

Met Ile Pro Ile Met Ser Ser Ala Pro Leu Ser Thr Ala Thr Pro Ile Ser Leu Pro Val Arg Lys Ile Pro Gly Ser Tyr Gly Phe Pro Leu Leu 25 Gly Pro Leu Trp Asp Arg Leu Asp Tyr Asn Trp Phe Gln Lys Leu Pro 40 Asp Phe Phe Ser Lys Arg Val Glu Lys Tyr Asn Ser Thr Val Phe Arg Thr Asn Val Pro Pro Cys Phe Pro Phe Phe Leu Gly Val Asn Pro Asn Val Val Ala Val Leu Asp Val Lys Ser Phe Ala His Leu Phe Asp Met Glu Ile Val Glu Lys Ala Asn Val Leu Val Gly Asp Phe Met Pro Ser 90 105 Val Val Tyr Thr Gly Asp Met Arg Val Cys Ala Tyr Leu Asp Thr Ser 120 Glu Pro Lys His Thr Gln Ile Lys Asn Phe Ser Leu Asp Ile Leu Lys Arg Ser Ser Lys Thr Trp Val Pro Thr Leu Val Lys Glu Leu Asp Thr Leu Phe Gly Thr Phe Glu Ser Asp Leu Ser Lys Ser Lys Ser Ala Ser 170 Leu Leu Pro Ala Leu Gln Lys Phe Leu Phe Asn Phe Phe Ser Leu Thr Phe Leu Gly Ala Asp Pro Ser Ala Ser Pro Glu Ile Ala Asn Ser Gly 185 200 Phe Ala Tyr Leu Asp Ala Trp Leu Ala Ile Gln Leu Ala Pro Thr Val Ser Ile Gly Val Leu Gln Pro Leu Glu Glu Ile Phe Val His Ser Phe Ser Tyr Pro Tyr Phe Leu Val Arg Gly Gly Tyr Glu Lys Leu Ile Lys 250

```
Phe Val Lys Ser Glu Ala Lys Glu Val Leu Thr Arg Ala Gln Thr Asp
            260
                                265
Phe Gln Leu Thr Glu Gln Glu Ala Ile His Asn Leu Leu Phe Ile Leu
       275
                            280
                                                285
Gly Phe Asn Ala Phe Gly Gly Phe Thr Ile Phe Leu Pro Thr Leu Leu
                        295
                                            300
Gly Asn Leu Gly Asp Glu Lys Asn Ala Glu Met Gln Glu Lys Leu Arg
                    310
                                        315
Lys Glu Val Arg Glu Lys Val Gly Thr Asn Gln Glu Asn Leu Ser Phe
               325
                                   330
Glu Ser Val Lys Glu Met Glu Leu Val Gln Ser Phe Val Tyr Glu Ser
            340
                                345
                                                    350
Leu Arg Leu Ser Pro Pro Val Pro Ser Gln Tyr Ala Arg Ala Arg Lys
       355
                            360
                                                365
Asp Phe Met Leu Ser Ser His Asp Ser Val Tyr Glu Ile Lys Lys Gly
   370
                        375
                                            380
Glu Leu Leu Cys Gly Tyr Gln Pro Leu Val Met Lys Asp Pro Lys Val
                    390
                                        395
                                                             400
Phe Asp Glu Pro Glu Lys Phe Met Leu Glu Arg Phe Thr Lys Glu Lys
                405
                                    410
                                                         415
Gly Lys Glu Leu Leu Asn Tyr Leu Phe Trp Ser Asn Gly Pro Gln Thr
            420
                                425
                                                    430
Gly Ser Pro Thr Glu Ser Asn Lys Gln Cys Ala Ala Lys Asp Ala Val
       435
                            440
Thr Leu Thr Ala Ser Leu Ile Val Ala Tyr Ile Phe Gln Lys Tyr Asp
                        455
                                            460
Ser Val Ser Phe Ser Ser Gly Ser Leu Thr Ser Val Lys Lys Ala Cys
<210> 5
<211> 517
<212> PRT
<213> Arabidopsis
Met Ala Ser Ile Ser Thr Pro Phe Pro Ile Ser Leu His Pro Lys Thr
                                   10
Val Arg Ser Lys Pro Leu Lys Phe Arg Leu Val Thr Arg Pro Ile Lys
            20
                               25
Ala Ser Gly Ser Glu Thr Pro Asp Leu Thr Val Ala Thr Arg Thr Gly
       35
                         40
Ser Lys Asp Leu Pro Ile Arg Asn Ile Pro Gly Asn Tyr Gly Leu Pro
                        55
                                            60
Ile Val Gly Pro Ile Lys Asp Arg Trp Asp Tyr Phe Tyr Asp Gln Gly
                    70
Ala Glu Glu Phe Phe Lys Ser Arg Ile Arg Lys Tyr Asn Ser Thr Val
                85
                                    90
                                                         95
Tyr Arg Val Asn Met Pro Pro Gly Ala Phe Ile Ala Glu Asn Pro Gln
            100
                                105
                                                     110
Val Val Ala Leu Leu Asp Gly Lys Ser Phe Pro Val Leu Phe Asp Val
        115
                            120
                                                125
Asp Lys Val Glu Lys Lys Asp Leu Phe Thr Gly Thr Tyr Met Pro Ser
                        135
                                            140
Thr Glu Leu Thr Gly Gly Tyr Arg Ile Leu Ser Tyr Leu Asp Pro Ser
```

60 120

```
Glu Pro Lys His Glu Lys Leu Lys Asn Leu Leu Phe Phe Leu Lys
                                     170
 Ser Ser Arg Asn Arg Ile Phe Pro Glu Phe Gln Ala Thr Tyr Ser Glu
 Leu Phe Asp Ser Leu Glu Lys Glu Ala Phe Pro Leu Arg Glu Ser Gly
                                 185
                             200
 Phe Arg Arg Phe Gln Arg Arg Asn Arg Leu Leu Phe Leu Gly Ser Ser
 Phe Leu Arg Asp Glu Ser Arg Arg Tyr Lys Leu Lys Ala Asp Ala Pro
 Gly Leu Ile Thr Lys Trp Val Leu Phe Asn Leu His Pro Leu Leu Ser
                                     250
 Thr Gly Leu Pro Arg Val Ile Glu Glu Pro Leu Ile His Thr Phe Ser
                                 265
 Leu Pro Pro Ala Leu Val Lys Ser Asp Tyr Gln Arg Leu Tyr Glu Phe
                             280
 Leu Arg Ile Arg Gly Glu Ile Leu Val Glu Ala Asp Lys Leu Gly Ile
 Ser Arg Glu Glu Ala Thr His Asn Leu Leu Phe Ala Thr Ser Phe Asn
                                             300
Thr Trp Gly Gly Met Lys Ile Leu Phe Pro Asn Met Val Lys Arg Ile
                                    330
Gly Pro Gly Gly His Gln Val His Asn Arg Leu Ala Glu Glu Ile Arg
                                345
Ser Val Ile Lys Ser Asn Gly Gly Glu Leu Thr Met Gly Ala Ile Glu
                            360
Lys Met Glu Leu Thr Lys Ser Val Val Tyr Glu Cys Leu Arg Phe Glu
Pro Pro Val Thr Ala Gln Tyr Gly Arg Ala Lys Lys Asp Leu Val Ile
                                       395
Glu Ser His Asp Ala Ala Phe Lys Val Lys Ala Gly Glu Met Leu Tyr
Gly Tyr Gln Pro Leu Ala Thr Arg Asp Pro Lys Ile Phe Asp Arg Ala
                                    410
                                425
Asp Glu Phe Val Pro Glu Arg Phe Val Gly Glu Gly Glu Lys Leu
                            440
Leu Arg His Val Leu Trp Ser Asn Gly Pro Glu Thr Glu Thr Pro Thr
                        455
Val Gly Asn Lys Gln Cys Ala Gly Lys Asp Phe Val Val Leu Val Ala
Arg Leu Phe Val Ile Glu Ile Phe Arg Arg Tyr Asp Ser Phe Asp Ile
                                  490
Glu Val Gly Thr Ser Pro Leu Gly Ser Ser Val Asn Phe Ser Ser Leu
                               505
Arg Lys Ala Ser Phe
                                                    510
       515
```

<210> 6

<211> 1901

<212> DNA

<213> Lycopersicon eculentum

<400> 6

ttgataatga tattcagett accacecaac geeetcaace ttgactagee etceetcaaa etcteetttt teaccaatet caetetecaa atagtatttt etaggtttea tagtgataag

- 6 -

```
cacatccaca caagagctat gaaatttctg aaacaaacgt tacaatagac acacggcgta
                                                                                           180
tttatttccc tcaagcaatc gtatccctt agcagatcga aaaaaaaaga agataccgat aatgaattct gctcctctat caactccggc gccggttact ctgcccgtcc gttcgattcc gggcagctac ggattgccgt tagtagggcc aatcgcggat cgattagact acttctggtt
                                                                                           240
                                                                                           300
                                                                                           360
ccaaaaaccg gaaaacttct tcaccaagag aatggaaaag cacaagagca cggttttcag
                                                                                           420
aacgaacgtg ccgccgtgtt ttccgttttt cggtagtgtg aatccaaatg tggtggcggt
                                                                                           480
tttagatgtg aaatcgtttt cgcatctgtt tgatatggag attgtagaga aagctaatgt gcttgttggg gatttcatgc ccagtgttgt ttatactgga gatatgcgtg tttgtgctta
                                                                                           540
                                                                                           600
tctcgatact tctgaaccta aacatgccca gattaagaat ttttcacagg atattctaaa
                                                                                           660
aagaggetea aaaacatggg tgeetacaet aettaaagag ettgatacaa tgtttacaae
                                                                                           720
ttttgaagca gatctttcaa aatccaatac agcttctctt cttcctgcac tccaaaaatt cctctttaac ttcttctccc tcactatcct cggcgctgat ccatctgtct ccccagaaat
                                                                                           780
                                                                                           840
cgccaattct ggctacatct ttcttgattc atggctagct attcaattag cacctactgt
                                                                                           900
cagcattggt gtccttcaac ctcttgaaga aattttagtc cattcttttg catacccttt
                                                                                           960
ttttcttgtc aaaggtaatt atgaaaaact cgttcaattc gtcaaaaatg aagccaagga
                                                                                          1020
agttttaagc agggcacaaa cagagtttca actcacagaa caagaagcca ttcataacct
                                                                                          1080
tttgttcatt cttgggttca atgcttttgg tggcttctcc attttcttgc caactctttt
                                                                                          1140
gggaaatctt ggagatgaga aaacgcgga tatgcaagag aaactgagaa aagaagtgag
                                                                                          1200
agacaaagtc ggcgtaaatc cagaaaactt gagttttgaa agtgttaaag aaatggaact
                                                                                          1260
tgttcagtct tttgtttatg aaacacttag gcttagtcca ccagtaccaa gtcaatatgc aagagcaaga aaagatttta aactgagttc acatgattca gtttacgaaa tcaagaaagg
                                                                                          1320
                                                                                          1380
ggagettett tgtggttate ageetttagt tatgaaagat ccaaaggtgt ttgatgaace
                                                                                          1440
tgaaaagttt gtgttggaga gatttacaaa ggaaaaaggg aaagaattgc tgaattattt
                                                                                          1500
gttttggtct aatggtccac aaactgggag acctactgaa tcaaacaagc aatgtgctgc taaagatatg gttactctaa ctgcttcttt gattgtggct tacattttcc aaaagtacga ttctgtgagt ttctcatctg gttcactcac atctgtgaaa aaggctagtt gagttctggg
                                                                                          1560
                                                                                          1620
                                                                                          1680
ctgtgtagct gcaaaatcta agcttgctag acatgttgat cttaagtctc ggtggctggc
ttttccttcc tccccctatt gtgtaatata atttcacaaa taaaagttca acatgagtgt
                                                                                          1800
acttattttc tgttctgact tatgactatt gtctatcctc ttacaaatta agttggaaca
                                                                                          1860
tgaatctaag ttataattac aaaaaaaaaa aaaaaaaaa a
                                                                                          1901
```

<210> 7 <211> 476 <212> PRT

<213> Lycopersicon eculentum

<220>

<223> Synthetic Oligonucleotide

<400> 7 Met Asn Ser Ala Pro Leu Ser Thr Pro Ala Pro Val Thr Leu Pro Val Arg Ser Ile Pro Gly Ser Tyr Gly Leu Pro Leu Val Gly Pro Ile Ala 25 Asp Arg Leu Asp Tyr Phe Trp Phe Gln Lys Pro Glu Asn Phe Phe Thr Lys Arg Met Glu Lys His Lys Ser Thr Val Phe Arg Thr Asn Val Pro Pro Val Phe Pro Phe Gly Ser Val Asn Pro Asn Val Val Ala Val 75 70 Leu Asp Val Lys Ser Phe Ser His Leu Phe Asp Met Glu Ile Val Glu 90 -85 95 Lys Ala Asn Val Leu Val Gly Asp Phe Met Pro Ser Val Val Tyr Thr 105 110 Gly Asp Met Arg Val Cys Ala Tyr Leu Asp Thr Ser Glu Pro Lys His

```
115
                              120
  Ala Gln Ile Lys Asn Phe Ser Gln Asp Ile Leu Lys Arg Gly Ser Lys
  Thr Trp Val Pro Thr Leu Leu Lys Glu Leu Asp Thr Met Phe Thr Thr
  Phe Glu Ala Asp Leu Ser Lys Ser Asn Thr Ala Ser Leu Leu Pro Ala
                                          155
 Leu Gln Lys Phe Leu Phe Asn Phe Phe Ser Leu Thr Ile Leu Gly Ala
                                  185
 Asp Pro Ser Val Ser Pro Glu Ile Ala Asn Ser Gly Tyr Ile Phe Leu
                              200
 Asp Ser Trp Leu Ala Ile Gln Leu Ala Pro Thr Val Ser Ile Gly Val
 Leu Gln Pro Leu Glu Glu Ile Leu Val His Ser Phe Ala Tyr Pro Phe
 Phe Leu Val Lys Gly Asn Tyr Glu Lys Leu Val Gln Phe Val Lys Asn
 Glu Ala Lys Glu Val Leu Ser Arg Ala Gln Thr Glu Phe Gln Leu Thr
                                      250
                                  265
 Glu Gln Glu Ala Ile His Asn Leu Leu Phe Ile Leu Gly Phe Asn Ala
 Phe Gly Gly Phe Ser Ile Phe Leu Pro Thr Leu Leu Gly Asn Leu Gly
 Asp Glu Lys Asn Ala Asp Met Gln Glu Lys Leu Arg Lys Glu Val Arg
 Asp Lys Val Gly Val Asn Pro Glu Asn Leu Ser Phe Glu Ser Val Lys
 Glu Met Glu Leu Val Gln Ser Phe Val Tyr Glu Thr Leu Arg Leu Ser
                                     330
 Pro Pro Val Pro Ser Gln Tyr Ala Arg Ala Arg Lys Asp Phe Lys Leu
 Ser Ser His Asp Ser Val Tyr Glu Ile Lys Lys Gly Glu Leu Leu Cys
Gly Tyr Gln Pro Leu Val Met Lys Asp Pro Lys Val Phe Asp Glu Pro
Glu Lys Phe Val Leu Glu Arg Phe Thr Lys Glu Lys Gly Lys Glu Leu
Leu Asn Tyr Leu Phe Trp Ser Asn Gly Pro Gln Thr Gly Arg Pro Thr
Glu Ser Asn Lys Gln Cys Ala Ala Lys Asp Met Val Thr Leu Thr Ala
                                425
Ser Leu Ile Val Ala Tyr Ile Phe Gln Lys Tyr Asp Ser Val Ser Phe
Ser Ser Gly Ser Leu Thr Ser Val Lys Lys Ala Ser
                                             460
                    470
<210> 8
<211> 1712
<212> DNA
<213> Cucumis sativus
<400> 8
ctcctcttct ctctaccgga aaagttcaac caacctctcc cctccctttc aaattcatac
```

60

120

teccacatet etcaaatate gatgeacete accatgacet eeggeggaat geettecata

cetteateta ttteccecce geeggtetet ttacetetea gaaatateee eggeagetae

```
ggtttgccgc tgctgggatc catcggtgac cgactcgatt actactggtt tcaaggaccc
                                                                                  240
                                                                                  300
gataagttct tcaggactcg tatggagaag aatcgaagca cggttttcag aaccaatgtt
cctccgtctt tccctttcat ctccgctgat cccagagtag ttgcggttct ggattgcaag
                                                                                  360
tectttgege atetatttga catggaaate gtggagaaaa ataatgttet ggteggtgat tteatgeeca geateagttt caeeggaaae atgagagtgt gtgettattt ggataceteg
                                                                                  420
                                                                                  480
qaacccaatc actccaaqqt aaagaacttc attacagaca ttctacggcg gagctcaagg
                                                                                  540
atatggatat cggagttgga atccaaccta tcgacgatgt gggacggcat tgaattagaa
                                                                                  600
atggcaaaga acaaacaatc gggtttcaga aactttctcc aaccagctct tttcaatttc ttctccaaaa ctctcgccgg cgccgacact gcaaaatcac cggaagtggc aaaatccgga
                                                                                  660
                                                                                  720
tacatagacy tcataatcty gctcggcctc cagctcgttc ccaccatcca catcggcgtt
                                                                                  780
ctccagcccc ttgaagaaat attccttcac tctttccgct tacccttttt ccccattgct
                                                                                  840
tecegetace aaagacteta egatteette caaaaagaag gggcagaagt tgttgagega ggtgtcaceg agttegggtt aacaaaggaa gaageeatte acaateteat etteaceatg
                                                                                  900
                                                                                  960
ggcttcaacg cctatggtgg cttcagtctc ttcttccctg ttctacttga ccgtatactc
                                                                                 1020
aatgataaaa ccggtctaca acaaagaatc ctcaaggaag tcaggtcaaa atccagctct
                                                                                 1080
ggtctgacct tcgagtcagt caaggagatg gacctcatct actccatcgt ctacgagaca ctccggcttg acccccagt tccctccaa tatgccagag ccagaagatt tcaagctaac
                                                                                 1140
                                                                                 1200
ttcctacaat tcaacgtata acatcaagaa aggggaactg ctttgtgggt atcagccgct
                                                                                 1260
ggtgatgaga gatccagagg tgttcgatga accggaggcg tttaatccag accgattccg
                                                                                . 1320
gggagagaaa ggggcagcgc tgctggatta tttattttgg tcgaatgggc cgcaaacggg
                                                                                 1380
gacaccgagc gagaagaaca aacagtgtgc tggggaaggat ttagtggtgc tgacgggggt ggtgtttgtg gcttacatat ttaaaaggta tgattcgatt gctggggaag gaggttccat
                                                                                 1440
                                                                                 1500
tacagetttt caaagggeta attgaaaagt gaaagtgaat tatatataca ttatgtateg
                                                                                 1560
tttattgaga gatgcacttt tcttgttcaa tggcttcttt cttatgtatg agtttggagc
                                                                                 1620
caaaatggga agatttggaa tgaataaatc aataaaatca agatttactt taccttgtaa
                                                                                 1680
aaaaaaaaa aaaaaaaaa aaaaaaaaaa aa
                                                                                 1712
<210> 9
<211> 1813
<212> DNA
<213> Cucumis sativus
<400> 9
ccattetett ccaacgtgaa gataagaaac ettttgttta etttttgtae gateacaggt cacagcaatg gettetteet eeettgaact teeteteaaa eccatteeeg gtggetatgg
                                                                                    60
                                                                                   120
                                                                                   180
cttccccttc ctcggtccca tcaaagaccg ttacgattac ttctatttcc aaggtagaga
                                                                                   240
cgaattette cgtteecgaa ttaccaaata caactecace gtetteeacg ecaacatgee
accgggcccc ttcatctcct ccgattccag agtcgttgtc ctcctcgatg ccctcagttt
                                                                                   300
teccatecte ttegacacca ecaaagtega gaaacgeaac attetegacg gaacttacat
                                                                                   360
geoctecttg teetteaceg geggeatteg cacetgtget tatttggace categgaaac
                                                                                   420
agagcacact gttctcaaac gcctctttct ctcctttctc gcttctcacc atgacaggtt
                                                                                   480
                                                                                   540
catecetetg tttegaaget eettgtetga gatgtttgtt aagettgaag ataaactege
cgacaaaaat aagatcgctg atttcaactc gattagtgat gccgtgtcgt ttgattatgt
                                                                                   600
tttccgttta ttctccgatg gaacccctga ttcgacatta gctgctgatg gacctggaat gttcgattta tggcttgggc ttcaacttgc cccattggct tccattggcc ttcccaaaat
                                                                                   660
                                                                                   720
                                                                                   780
 tttctctgtt tttgaagate tcattattca caccattccc ctgcctttct tcccagtcaa
gagtcgttac aggaagcttt ataaagcgtt ttactcctcc tctggctcat ttctagacga
                                                                                   840
agcagagaaa caggggatag acagagagaa agcatgtcac aatttagtgt ttcttgctgg
                                                                                   900
attcaacgca tacgggggaa tgaaagtcct ttttcccact atactgaaat gggtcggcac
                                                                                   960
                                                                                  1020
 cggtggcgag gatctccacc gtaaactggc ggaggaagtg aggacaaccg tgaaggaaga
 agggggactg actttctccg ccttggagaa aatgagtctg ctgaagtcgg tcgtgtacga
                                                                                  1080
 agctctgagg atcgaaccgc cggtgccgtt ccagtacggg aaagcgaagg aggatatcgt
                                                                                  1140
gattcagagc cacgattctt gtttcaagat caaaaaaggg gagacgattt ttggttatca
                                                                                  1200
 gccgtttgct actaaagatc cgaagatttt taaggactcg gagaagttcg tgggcgatag
                                                                                  1260
```

1320

gttcgtggga gaagaagggg agaagctttt gaagtatgtt tactggtcaa atgagcggga

1440

1500

1560

1620 1680

```
gacggtggag ccgacggcgg agaacaagca gtgtccgggg aagaatctgg tggtgatgat
  gggcaggatt attgtggtgg aattetteet eegttatgat aegtteaetg tggaegtege
  agatttggcg ctgggcccgg cggtgaagtt caagtccttg accagagcca ccgcttcggt
  ttagaagcta atgacaaaat tagtttttaa tcattttaca gaataaagga agggtgttcg
  tgattggttg atgcagggaa gctgtggtgg acttgagcac acatgacaca tcattgattt
  gttggggtta taaagtaatt teetegtgat eegegtegtt tttaatttga gateteattg
  tgtgttgtaa cccgcgggtg atcttatttt tatagtttct tttttctcaa ctatgctcca
  <210> 10
  <211> 477
  <212> PRT
  <213> Cucumis sativus
 <220>
 <223> Synthetic Oligonucleotide
 <400> 10
 Met Ala Ser Ser Pro Glu Leu Pro Leu Lys Pro Ile Pro Gly Gly
 Tyr Gly Phe Pro Phe Leu Gly Pro Ile Lys Asp Arg Tyr Asp Tyr Phe
 Tyr Phe Gln Gly Arg Asp Glu Phe Phe Arg Ser Arg Ile Thr Lys Tyr
 Asn Ser Thr Val Phe His Ala Asn Met Pro Pro Gly Pro Phe Ile Ser
 Ser Asp Ser Arg Val Val Leu Leu Asp Ala Leu Ser Phe Pro Ile
 Leu Phe Asp Thr Thr Lys Val Glu Lys Arg Asn Ile Leu Asp Gly Thr
 Tyr Met Pro Ser Leu Ser Phe Thr Gly Gly Ile Arg Thr Cys Ala Tyr
Leu Asp Pro Ser Glu Thr Glu His Thr Val Leu Lys Arg Leu Phe Leu
Ser Phe Leu Ala Ser His His Asp Arg Phe Ile Pro Leu Phe Arg Ser
Ser Leu Ser Glu Met Phe Val Lys Leu Glu Asp Lys Leu Ala Asp Lys
Asn Lys Ile Ala Asp Phe Asn Ser Ile Ser Asp Ala Val Ser Phe Asp
Tyr Val Phe Arg Leu Phe Ser Asp Gly Thr Pro Asp Ser Thr Leu Ala
Ala Asp Gly Pro Gly Met Phe Asp Leu Trp Leu Gly Leu Gln Leu Ala
Pro Leu Ala Ser Ile Gly Leu Pro Lys Ile Phe Ser Val Phe Glu Asp
Leu Ile Ile His Thr Ile Pro Leu Pro Phe Phe Pro Val Lys Ser Arg
Tyr Arg Lys Leu Tyr Lys Ala Phe Tyr Ser Ser Gly Ser Phe Leu Asp
Glu Ala Glu Lys Gln Gly Ile Asp Arg Glu Lys Ala Cys His Asn Leu
Val Phe Leu Ala Gly Phe Asn Ala Tyr Gly Gly Met Lys Val Leu Phe
```

```
Pro Thr Ile Leu Lys Trp Val Gly Thr Gly Glu Asp Leu His Arg
                        295
                                            300
Lys Leu Ala Glu Glu Val Arg Thr Thr Val Lys Glu Glu Gly Gly Leu
305
                    310
                                        315
                                                             320
Thr Phe Ser Ala Leu Glu Lys Met Ser Leu Leu Lys Ser Val Val Tyr
                325
                                    330
                                                         335
Glu Ala Leu Arg Ile Glu Pro Pro Val Pro Phe Gln Tyr Gly Lys Ala
           340
                                345
Lys Glu Asp Ile Val Ile Gln Ser His Asp Ser Cys Phe Lys Ile Lys
       355
                            360
Lys Gly Glu Thr Ile Phe Gly Tyr Gln Pro Phe Ala Thr Lys Asp Pro
                        375
                                            380
Lys Ile Phe Lys Glu Ser Glu Lys Phe Val Gly Asp Arg Phe Val Gly
385
                    390
                                        395
Glu Glu Gly Glu Lys Leu Leu Lys Tyr Val Tyr Trp Ser Asn Glu Arg
                405
                                    410
                                                         415
Glu Thr Val Glu Pro Thr Ala Glu Asn Lys Gln Cys Pro Gly Lys Asn
                                425
                                                     430
            420
Leu Val Val Met Met Gly Arg Ile Ile Val Val Glu Phe Phe Leu Arg
        435
                            440
                                                 445
Tyr Asp Thr Phe Thr Val Asp Val Ala Asp Leu Ala Leu Gly Pro Ala
                        455
Val Lys Phe Lys Ser Leu Thr Arg Ala Thr Ala Ser Val
                    470
```

<210> 11 <211> 482 <212> PRT

BNSDOCID: <WO__0060575A2 1 >

<213> Banana

<400> 11 Met Ala Met Met Trp Ser Ser Ala Ser Ala Thr Ala Val Thr Thr Leu Pro Thr Arg Pro Ile Pro Gly Ser Tyr Gly Pro Pro Leu Val Gly Pro Leu Lys Asp Arg Leu Asp Tyr Phe Trp Phe Gln Gly Pro Glu Thr Phe 40 Phe Arg Ser Arg Met Ala Thr His Lys Ser Thr Val Phe Arg Thr Asn 50 55 Met Pro Pro Thr Phe Pro Phe Phe Val Gly Val Asp Pro Arg Val Val 75 Thr Val Leu Asp Cys Thr Ser Phe Ser Ala Leu Phe Asp Leu Gln Val 85 Val Glu Lys Lys Asn Ile Leu Ile Gly Asp Tyr Met Pro Ser Leu Ser . 100 105 110 Phe Thr Gly Asp Thr Arg Val Val Val Tyr Leu Asp Pro Ser Glu Pro 115 120 125 Asp His Ala Arg Val Lys Ser Phe Cys Leu Gln Leu Leu Arg Arg Gly 135 140 Ala Lys Thr Trp Val Ser Ser Phe Leu Ser Asn Leu Asp Val Met Leu 150 155 Ala Thr Ile Glu Gln Gly Ile Ala Lys Asp Gly Ser Ala Gly Leu Phe 170 165 Gly Pro Leu Gln Lys Cys Ile Phe Ala Phe Leu Cys Lys Ser Ile Ile 180 185

```
Gly Ala Asp Pro Ser Val Ser Pro Asp Val Gly Glu Asn Gly Phe Val
 Met Leu Asp Lys Trp Leu Ala Leu Gln Leu Leu Pro Thr Val Lys Val
 Gly Ala Ile Pro Gln Pro Leu Glu Glu Ile Leu Leu His Ser Phe Pro
 Leu Pro Phe Phe Leu Val Ser Arg Asp Tyr Arg Lys Leu Tyr Glu Phe
                                     250
 Val Glu Lys Gln Gly Gln Glu Val Val Arg Arg Ala Glu Thr Glu His
                                 265
 Gly Leu Ser Lys His Asp Ala Ile Asn Asn Ile Leu Phe Val Leu Gly
Phe Asn Ala Phe Gly Gly Phe Ser Val Phe Phe Pro Thr Leu Leu Thr
Thr Ile Gly Arg Asp Lys Thr Gly Leu Arg Glu Lys Leu Lys Asp Glu
Val Arg Arg Val Met Lys Ser Arg Gly Glu Lys Arg Pro Ser Phe Glu
Thr Val Arg Glu Met Glu Leu Val Arg Ser Thr Val Tyr Glu Val Leu
                                    330
Arg Leu Asn Pro Pro Val Pro Leu Gln Gly Arg Ala Arg Thr Asp Phe
Thr Leu Asn Ser His Asp Ala Ala Phe Lys Val Glu Lys Gly Glu Leu
Leu Cys Gly Tyr Gln Pro Leu Val Met Arg Asp Pro Ala Val Phe Asp
Asp Pro Glu Thr Phe Ala Pro Glu Arg Phe Met Gly Ser Gly Lys Glu
Leu Leu Lys Tyr Val Phe Trp Ser Asn Gly Pro Glu Thr Gly Thr Pro
Thr Pro Ala Asn Lys Gln Cys Ala Ala Lys Asp Tyr Val Val Glu Thr
Ala Cys Leu Leu Met Ala Glu Ile Phe Tyr Arg Tyr Asp Glu Phe Val
Cys Ala Asp Asp Ala Ile Ser Val Thr Lys Leu Asp Arg Ala Arg Glu
                                        475
Trp Glu
```

28

- 12 -

<210> 12

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Oligonucleotide

<400> 12

cggttcctct gcgcctctct cgccggcg

<210> 13

<211> 26

<212> DNA

<213> Artificial Sequence

<220> <223> Synthetic Oligonucleotide	
<400> 13 gcggaaccgg aggactaaaa cgcagc	26
<210> 14 <211> 27 <212> DNA <213> Artificial Sequence	
<220> <223> Synthetic Oligonucleotide	
<400> 14 ccatcctaat acgactcact ataggge	27
<210> 15 <211> 25 <212> DNA <213> Artificial Sequence	•
<220> <223> Synthetic Oligonucleotide	
<400> 15 cttggcgtag ttcctcagcc tcttg	25
<210> 16 <211> 23 <212> DNA <213> Artificial Sequence	
<220> <223> Synthetic Oligonucleotide	
<400> 16 actcactata gggctcgagc ggc	23
<210> 17 <211> 32 <212> DNA <213> Artificial Sequence	
<220> <223> Synthetic Oligonucleotide	
<400> 17 cgggatccat gttgttgaga acgatggcgg cg	32
<210> 18 <211> 21 <212> DNA <213> Artificial Sequence	

```
<220>
  <223> Synthetic Oligonucleotide
  <400> 18
 caatctccgg cgttctcgtc g
                                                                            21
 <210> 19
  <211> 17
  <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic Oligonucleotide
 <221> misc_feature
 <222> (1)...(17)
 <223> n = A,T,C or G
 <400> 19
 atnccnggnw sntaygg
                                                                           17
 <210> 20
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic Oligonucleotide
 <221> misc_feature
 <222> (1) ... (17)
 <223> n = A,T,C or G
<400> 20
carconytng argarat
                                                                           17
<210> 21
<211> 17
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic Oligonucleotide
<221> misc_feature
<222> (1)...(17)
<223> n = A,T,C or G
<400> 21
atytcytcna rnggytg
                                                                          17
<210> 22
<211> 20
<212> DNA
<213> Artificial Sequence
```

<220> <223> Synthetic Oligonucleotide	
<221> misc_feature <222> (1)(20) <223> n = A,T,C or G	
<400> 22 ggnttyaayg cntwyggngg	20
<210> 23 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> Synthetic Oligonucleotide	
<221> misc_feature <222> (1)(20) <223> n = A,T,C or G	
<400> 23 ccnccrsang crttraance	20
<210> 24 <211> 18 <212> DNA <213> Artificial Sequence	;
·	
<220> <223> Synthetic Oligonucleotide	
	÷
<223> Synthetic Oligonucleotide <221> misc_feature <222> (1)(18)	18
<223> Synthetic Oligonucleotide <221> misc_feature <222> (1)(18) <223> n = A,T,C or G <400> 24	18
<223> Synthetic Oligonucleotide <221> misc_feature <222> (1)(18) <223> n = A,T,C or G <400> 24 taycarccny tngtnatg <210> 25 <211> 18 <212> DNA	18
<221> misc_feature <222> (1)(18) <223> n = A,T,C or G <400> 24 taycarccny tngtnatg <210> 25 <211> 18 <212> DNA <213> Artificial Sequence <220>	18

```
<210> 26
<211> 17
  <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic Oligonucleotide
 <221> misc_feature
 <222> (1)...(17)
<223> n = A,T,C or G
 <400> 26
 gtnttygayg anccnga
                                                                                17
 <210> 27
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic Oligonucleotide
<221> misc_feature
 <222> (1)...(17)
 <223> n = A,T,C or G
 <400> 27
tcnggntcrt craanac
                                                                               17
<210> 28
<211> 17
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic Oligonucleotide
<221> misc_feature
<222> (1) ... (17)
<223> n = A,T,C \text{ or } G
<400> 28
cengtytwng gneertt
                                                                               17
<210> 29
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic Oligonucleotide
<221> misc_feature
```

<222> (1)(22) <223> n = A,T,C or G	
<400> 29 cyttngcngc rcaytgyttr tt	22
<210> 30 <211> 31 <212> DNA <213> Artificial Sequence	
<220> <223> Synthetic Oligonucleotide	
<400> 30 cggtggagat cctcgccacc ggtgccgacc c	31
<210> 31 <211> 34 <212> DNA <213> Artificial Sequence	
<220> <223> Synthetic Oligonucleotide	
<400> 31 cttccttcac ggttgtcctc acttcctccg ccag	34
<210> 32 <211> 30 <212> DNA <213> Artificial Sequence	
<220> <223> Synthetic Oligonucleotide	
<400> 32 tccagcagcg ctgccccttt ctctccccgg	30
<210> 33 <211> 31 <212> DNA <213> Artificial Sequence	
<220> <223> Synthetic Oligonucleotide	
<400> 33 cactgtttgt tetteteget eggtgteece g	31
<210> 34 <211> 31 <212> DNA <213> Artificial Sequence	

PCT/US00/05311

<220> <223> Synthetic Oligonucleotide			
<400> 34 gggtcggcac cggtggcgag gatctccacc	: a		
<210> 35 <211> 34	3		31
<212> DNA <213> Artificial Sequence			1.5
<220> <223> Synthetic Oligonucleotide			
<400> 35 ctggcggagg aagtgaggac aaccgtgaag	gaag		• •
<210> 36 <211> 29 <212> DNA <213> Artificial Sequence			34
<220> <223> Synthetic Oligonucleotide			
<400> 36 ccggggagag aaaggggcag cgctgctgg			20
<210> 37 <211> 31 <212> DNA <213> Artificial Sequence			29
<220> <223> Synthetic Oligonucleotide			
<400> 37 cggggacacc gagcgagaag aacaaacagt	g	elegy.	31
<210> 38 <211> 34 <212> DNA <213> Artificial Sequence			
<220> <223> Synthetic Oligonucleotide			(i)
<400> 38 gacttggtac tggtggacta agcctaagtg (tttc		34
<210> 39 <211> 30 <212> DNA			34
< 213 \ \ None is in the control of the co		*	

<220> <223>	Synthetic Oligonucleotide		
<400> ggctg	39 ataac cacaaagaag ctcccctttc	3	0
<210><211><212><213>	34		
<220> <223>	Synthetic Oligonucleotide		
<400> gaaac	40 actta ggcttagtcc accagtacca	agtc 3	4
<210><211><211><212><213>	30		
<220> <223>	Synthetic Oligonucleotide		
<400> gaaag	41 gggag cttctttgtg gttatcagcc	** *** *** *** *** *** *** *** *** ***	30
<210><211><211><212><213>	33		•
<220> <223>	Synthetic Oligonucleotide	. ω "	٠.
<400> cggga	42 Atccat ggcttcttcc tcccctgaac	ttc	33 ့
<210><211><211><212><213>	- 25		,•
<220> <223>	Synthetic Oligonucleotide		
<400> tgccg	- 43 gaccca tttcagtata gtggg		25
<210><211><211><212><213>	> 28		•

<220> <223> Synthetic Oligonucleotide <400> 44 ttcacaccat tccctgcct ttcttccc

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 31 August 2000 (31.08.2000)

PCT

(10) International Publication Number WO 00/50575 A3

- (51) International Patent Classification7: C12N 9/88, 15/82
- (21) International Application Number: PCT/US00/05311
- (22) International Filing Date: 25 February 2000 (25.02.2000)
- (25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data: 60/121,968

26 February 1999 (26.02.1999) US

- (71) Applicant: CALGENE LLC [US/US]; 1920 Fifth Street, Davis, CA 95616 (US).
- (72) Inventor: MATSUI, Kenji; 2-51, Nishiki-cho, Yamaguchi City, Yamaguchi 753-0068 (JP).

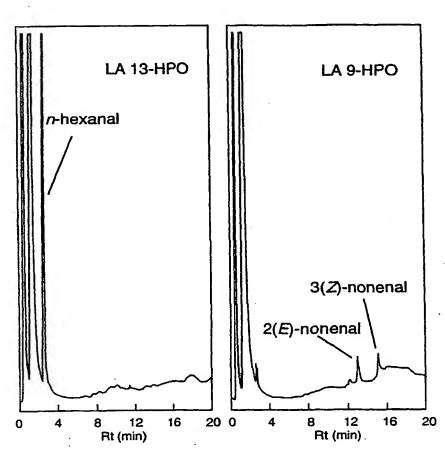
- (74) Agent: STIERWALT, Brian, K.; Monsanto, 800 N. Lindbergh Blvd., Mail Zone O3C, St. Louis, MO 63167 (US).
- (81) Designated States (national): CA, CN, JP, MX.
- (84) Designated States (regional): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published:

- with international search report
- (88) Date of publication of the international search report:
 13 September 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NUCLEIC ACID SEQUENCE OF A CUCUMBER (CUCUMIS SATIVUS) FATTY ACID 9-HYDROPEROXIDE LYASE



(57) Abstract: This invention 9-Hydroperoxide relates to Lyase or 9-HPO lyase enzymes. DNA constructs useful for the expression of a plant HPO lyase in a cell are described. Furthermore, DNA constructs useful for the antisense expression of a 9-HPO lyase in a plant cell are described. Such constructs will contain a DNA sequence encoding the 9-HPO lyase of interest under the control of regulatory elements capable of preferentially directing the expression of the 9-HPO lyase in plant tissue, when such a construct is expressed in a This invention trangenic plant. also relates to methods of using a DNA sequence encoding a 9-HPO lyase for the modification of the volatile aldehydes in plant tissues, as well as for methods of increasing disease resistance in a plant. 9-HPO Lyase sequences exemplified herein are obtained from Arabidopsis.

Internations plication No

PCT/US U0/05311 A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N9/88 C12N C12N15/82 According to International Patent Classification (IPC) or to both national classification and IPC **B. FIELDS SEARCHED** Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, BIOSIS, WPI Data, PAJ, MEDLINE, EMBASE, SCISEARCH, BIOTECHNOLOGY ABS, CHEM ABS Data, CAB Data, GENSEO, STRAND C. DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages Refevant to claim No. Υ MATSUI K ET AL: "SEPARATION OF 13 AND 9 1-11 HYDROPEROXIDE LYASE ACTIVITIES IN COTYLEDONS OF CUCUMBER SEEDLINGS" ZEITSCHRIFT FUER NATURFORSCHUNG SECTION C BIOSCIENCES, vol. 44, no. 9-10, 1989, pages 883-885, XP000949968 ISSN: 0341-0382 the whole document -/--Further documents are listed in the continuation of box C. X X Patent family members are listed in annex. Special categories of cited documents: T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "O" document referring to an oral disclosure, use, exhibition or other means document published prior to the international filing date but later than the priority date claimed in the art. "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 9 October 2000 23/10/2000 Name and mailing address of the ISA Authorized officer European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl. Fax: (+31-70) 340-3016 ALCONADA RODRIG.., A

Internation plication No
PCT/US U0/05311

		PC1/US 00/05311
C.(Continua	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to daim No.
Y	MATSUI ET AL: "Bell pepper fruit fatty acid hydroperoxide lyase is a cytochrome P450 (CYP74B)" FEBS LETTERS,NL,ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, vol. 394, no. 1, 23 September 1996 (1996-09-23), pages 21-24, XP002126014 ISSN: 0014-5793 the whole document	1-11
Y	BATE N J ET AL: "Molecular characterization of an Arabidopsis gene encoding hydroperoxide lyase, a cytochrome P-450 that is wound inducible" PLANT PHYSIOLOGY, US, AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS, ROCKVILLE, MD, vol. 117, no. 4, August 1998 (1998-08), pages 1393-1400, XP002128564 ISSN: 0032-0889 the whole document	1-11
A	NOORDERMEER M A ET AL: "Alfalfa contains substantial 9- hydroperoxide lyase activity and a 3Z:2E-enal isomerase" FEBS LETTERS,NL,ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, vol. 443, 25 January 1999 (1999-01-25), pages 201-204, XP002136824 ISSN: 0014-5793 the whole document	1-11
A	HATANAKA A ET AL: "EXPRESSION OF LIPOXYGENASE AND HYDROPEROXIDE LYASE ACTIVITIES IN TOMATO FRUITS" ZEITSCHRIFT FÜR NATURFORSCHUNG, vol. 47, no. 5/06, 1992, pages 369-374, XP002120138 the whole document	1-11
A	KIM I-S ET AL: "PARTIAL PURIFICATION AND PROPERTIES OF A HYDRO PER OXIDE LYASE FROM FRUITS OF PEAR PYRUS-COMMUNIS" JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY, vol. 29, no. 6, 1981, pages 1220-1225, XP000946489 ISSN: 0021-8561 the whole document -/	1-11

Internations pilication No

Category	: Citation of document, with indication, where appropriate, of the relevant passages		Datamatic
	· · · · · · · · · · · · · · · · · · ·		Relevant to claim No.
1	GARGOURI MOHAMED ET AL: "Biosynthesis and analysis of 3Z-nonenal." BIOTECHNOLOGY LETTERS,		1-11
	vol. 20, no. 1, January 1998 (1998-01), pages 23-26, XP000949969 ISSN: 0141-5492 the whole document		
i	FAUCONNIER ET AL: "Purification and characterization of tomato leaf (Lycopersicon esculentum Mill.) hydroperoxide lyase" JOURNAL OF AGRICULTURAL AND FOOD		1-11
	CHEMISTRY,US,AMERICAN CHEMICAL SOCIETY. WASHINGTON, vol. 45, no. 11, November 1997 (1997-11), pages 4232-4236, XP002126015 ISSN: 0021-8561 the whole document		
	SHIBATA ET AL: "Purification and properties of fatty acid hydroperoxide lyase from green bell pepper fruits"		1-11
ls.	PLANT AND CELL PHYSIOLOGY, vol. 36, no. 1, 1995, pages 147-156		ē .
: 3) ::	XP000949972 ISSN: 0032-0781 the whole document		
·	CROFT KEVAN P C ET AL: "Volatile products of the lipoxygenase pathway evolved from Phaseolus vulgaris (L.) leaves inoculated with Pseudomonas syringae pv. phaseolicola."	*	10
	PLANT PHYSIOLOGY (ROCKVILLE), vol. 101, no. 1, 1993, pages 13-24, XP000946545		(F)
	ISSN: 0032-0889 cited in the application the whole document		
The street of th	BATE NICHOLAS J ET AL: "Quantitative and qualitative differences in C6-volatile production from the lipoxygenase pathway in an alcohol dehydrogenase mutant of		11
	Arabidopsis thaliana." PHYSIOLOGIA PLANTARUM, vol. 104, no. 1, 1998, pages 97-104, XP000950009		•
	ISSN: 0031-9317 the whole document		•
(-	WO 00 00627 A (MATSUI KENJI ;CALGENE LLC (US)) 6 January 2000 (2000-01-06) claims 7-9; examples 8.9,10		1-11
į	- <u></u>		

Internation plication No PCT/US 00/05311

			/ 03311
C.(Continua Category °	ation) DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages		Relevant to claim No.
Calegoly			
T	HORNOSTAJ A R ET AL: "Purification of hydroperoxide lyase from cucumbers." FOOD CHEMISTRY, vol. 66, no. 2, August 1999 (1999-08), pages 173-180, XP000949974 ISSN: 0308-8146 figures 2,4,5; tables 1-3		1-11
Τ .	DATABASE GENEMBL 'Online! 17 April 2000 (2000-04-17) MATSUI ET AL.: "Cucumis sativa fatty acid hydroperoxide lyase mRNA, complete cds" XP002148901 Accession number AF229811 -& MATSUI ET AL.: "Fatty acid 9- and 13-hydroperoxide lyases from cucumber" FEBS LETTERS, vol. 481, no. 2, 15 September 2000 (2000-09-15), pages 183-188, XP000952042 abstract		1-11
Τ	DATABASE GENEMBL 'Online! 14 May 1999 (1999-05-14) SHELDRICK ET AL.: "Lycopersicum esculentum partial mRNA for hydroperoxide layse" XP002136823 Accession number AJ239065		1-11
T	MATSUI K ET AL: "MOLECULAR CLONING AND EXPRESSION OF ARABIDOPSIS FATTY ACID HYDROPEROXIDE LYASE" PLANT AND CELL PHYSIOLOGY, vol. 40, no. 5, May 1999 (1999-05), pages 477-481, XP000870044 ISSN: 0032-0781 the whole document		1-11
		•	
	· ·		

Info. ! ion on patent family members

Internation olication No PCT/US U0/05311

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date | P

Form PCT/ISA/210 (patent family annex) (July 1992)